MASTER OF SCIENCE

Sequence Dependence of the Folding of Kink-Turns in RNA

Ashraf, Saira

Award date:
2017

MSc Thesis

To achieve the academic degree
Master of Sciences (MSc) by Research

Sequence Dependence of the Folding of Kink-Turns in RNA

Saira Ashraf

University of Dundee

College of Life Sciences

Professor David Lilley's Lab

Dundee, December 2017

Table of Contents

List of Figures

List of Tables

Acknowledgements

I would like to express my deep gratitude to the many people without whom this goal would not have been achievable.

First and foremost, I would very much like to thank Professor David Lilley for allowing me to carry out this work in his lab and for his guidance and support throughout.

Secondly, I would like to thank all my colleagues, in particular Lin Huang, whose teachings, guidance and helpful discussions have led to the successful completion of this project work.

I would also like to thank Cancer Research UK for funding my research and studies allowing me to attain the higher degree of Masters by Research.

Finally, I would like to thank my family for their support and encouragement especially my parents for always being by my side and encouraging me to achieve my ambitions. Special thanks to my husband for being my strength and inspiration.

Declaration

I hereby declare that I am the author of this thesis. The work of which this thesis is an account of has been done solely by myself, unless otherwise indicated, and has not been previously accepted for another higher degree. Unless otherwise stated, I have consulted all the references cited in this thesis.

Signature: ……...

Date: ……

1. Abstract

The kink-turn in RNA is a secondary structural motif which produces a tight kink in the axis of the RNA helix. It exists in a dynamic equilibrium between the folded and the unfolded state in free solution. The population of the folded kink-turn species can be increased by the addition of metal ions which bind to and stabilise the kink-turn structure. Standard kink-turn RNAs differ in their ability to undergo metal-ion induced folding. For example, the ribosomal kink-turn, *Hm*Kt-7, readily folds into the kinked conformation in metal ions. The *Af*boxC/D snoRNA kink-turn, on the other hand, is unable to fold in metal ions and requires the binding of the L7Ae protein to fold. The reason for this variation lies in the kink-turn RNA sequence. Certain positions, namely G1b:A1n and A2b:G2n, are critical for metal ion-induced kink-turn folding and form the basis for the classification of a RNA sequence as a standard kink-turn motif. The 3b:3n position has also been identified as being very important for metal-ion induced folding of kink-turns with different base pairs providing varying degrees of folding ability. U3b:U3n allows moderate folding of the modified *Hm*Kt-7 kink-turn in the presence of metal ions. The *Af*boxC/D kink-turn, which has U3b:U3n, is thus expected to fold. However, this is not the case leading to this study of other positions in the kink-turn RNA sequence which could potentially affect its folding.

The -1b:-1n, 3b:3n and 4b:4n positions of the *Haloarcula marismortui* Kt-7 and *Archeoglobus fulgidus* box C/D kink-turn RNAs have been studied in this work. The effect of modifying these positions on the ability of the kink-turns RNA to fold on the addition of metal ions has been investigated. RNA folding was studied using the techniques of fluorescence resonance energy transfer (FRET) and gel electrophoresis. The systematic exchange of sequence elements between the *Hm*Kt-7 and *Af*boxC/D kink-turn RNAs has revealed the importance of sequence on kink-turn RNA folding. The -1b:-1n, 3b:3n and 4b:4n positions all have an additive effect on the folding ability of kink-turns with the most profound effect exerted by the -1b:-1n position. In general, *Hm*Kt-7 has selected sequence elements that are C-1b:G-1n, A3b:G3n and C4b:G4n, which promote the metal ion-induced folding of the

kink-turn. On the other hand, *Af*boxC/D, despite having the moderate folding U3b:U3n sequence element, has G4b:C4n and, in particular, G-1b:C-1n which inhibits the folding of the RNA into the kink-turn structure in metal ions alone.

Furthermore, the N^6-methylation of adenine is a naturally occurring modification in cellular RNA which is also prevalent in the kink-turn regions of RNA. Hence, the effect of this modification on kink-turn RNA folding in the presence of metal ions, as well as the L7Ae protein, was studied by FRET. Substitutions in the various positions of the *Hm*Kt-7 kink-turn RNA sequence have revealed a differential effect on kink-turn folding. While N^6-methyladenine in the 1n position completely prevents metal ion and protein-induced folding of the kink-turn, the 2b and 3b positions are more tolerant although folding is impaired to some extent. These results provide some valuable insights to the effect that this naturally occurring modification can have on kink-turn RNA folding should any be found.

On the whole, the folding characteristics of kink-turns as determined by their sequence, is in adoption to their function. For example, ribosomal and riboswitch kink-turns have generally selected sequence elements that allow folding into the kinked conformation in metal ions alone which may be in accordance to their biological duty requiring free formation of the kink-turn structure. On the other hand, box C/D snoRNA kink-turns have selected sequence elements that render them unable to fold in metal ions alone but requiring the binding of the L7Ae-type protein to fold which is actually the first step in the assembly of the biologically active box C/D snoRNP complex. Nevertheless, naturally occurring sequence modifications, such as N^6-adenine methylation, can also alter the folding capability of kink-turns. A recently found example suggests that the naturally occurring N^6-methyladenine modification in the 1n position of the box C/D kink-turn RNA sequence can prevent the binding of the 15.5k protein, and hence prevent the assembly of the box C/D snoRNP complex. Thus, overall, in addition to sequence determinants, modifications such as N^6-adenine methylation can modulate RNA structure and consequently regulate RNA function.

2. Introduction

2.1 RNA – In General

Ribonucleic acid (RNA) is a macromolecule found in all living cells. RNA is vital for cell viability and takes part in various cellular processes. First and foremost, RNA acts as the intermediate element in the path of genetic information transmission, known as the central dogma [1]. Deoxyribonucleic acid (DNA) which contains the genes is transcribed into messenger RNA (mRNA) in the nucleus of the cell, which is then translated into protein in the ribosome in the cytoplasm [2]. The RNA transcribed from DNA can also be non-coding [3] i.e. does not code for proteins. The most prominent example of non-coding RNA is ribosomal RNA (rRNA) which is a structural component of ribosomes [4], as well as transfer RNA (tRNA) which mediates the transfer of the appropriate amino acid to the growing polypeptide chain during translation [5].

Other non-coding RNAs have also been identified which play vital roles in cells including the regulation of gene expression. These include ribozymes which have enzymatic activity [6] catalysing various cellular processes such as splicing [7]; and riboswitches, found in untranslated regions (UTRs) of mRNA [8, 9], which can alter conformation in response to external, e.g. temperature [7], and internal factors, e.g. various metabolites, such as guanidine [10], to modulate expression of the adjacent gene. Other long non-coding (lnc) RNA also exist, for example MALAT1 which regulates alternative splicing [11, 12]. Noncoding small regulatory RNAs (sRNAs) are also prevalent and include small nuclear RNAs (snRNAs), which in association with proteins form snRNP complexes, that splice the pre-mRNA into the mature mRNA [3, 13, 14]; microRNAs (miRNAs) which bind to 3'UTRs of mRNAs repressing their translation [3, 15]; small interfering RNAs (siRNAs) which associate with proteins to form siRNP complexes that target the sequence-specific degradation of mRNA [16-18]; and finally small nucleolar RNAs (snoRNAs) [19], which can be divided into the box C/D and H/ACA families [20], which in association with proteins form snoRNP complexes that

target the site-specific 2'-*O*-ribose methylation [21] and pseudouridylation [22, 23] of rRNA, respectively.

Overall, RNA is vital for normal cell functioning as it not only partakes important roles in the gene expression pathway as a messenger, transfer and ribosomal RNA, but also regulates gene expression by the means of controlling splicing, translation, mRNA degradation and rRNA modification. Defects in RNA can lead to various human diseases including cancer[24, 25] and heart disease [26, 27]. The proper functioning of RNA is dependent on its ability to adopt the correct structure.

2.2 RNA Structure

RNA is a polymeric molecule made from building blocks called nucleotides. A nucleotide comprises of a five-carbon ribose sugar to which is attached a phosphate group and a nucleobase. There are four types of nucleobase in RNA – adenine (A), guanine (G), cytosine (C) and uracil (U). These nucleobases can be divided into two classes: purines (A and G) which have a double-ringed structure and pyrimidines (C and U) which have a single-ringed structure. The nucleotides are linked together in varying sequences to form chains of various lengths. In the chain the 5'-phosphate group of a nucleotide is coupled to the 3'-hydroxyl group of the ribose of the previous nucleotide (Fig. 1).

RNA normally exists as a single-stranded molecule [28]. Two single strands of RNA can hybridise to form double-stranded helical structures [29] which may result in a secondary structure that is important to the function of the RNA concerned. Single-stranded RNA molecules can also fold into various secondary structures which, again, will be important for their correct functioning. One example of such a secondary structure is the hairpin loop. This involves the folding of a single RNA molecule which complementarily base pairs with another section of the same strand resulting in a loop like structure. The function of the hairpin loop includes to act as a site to bind proteins, for example to promote the perinuclear localisation of mRNA of

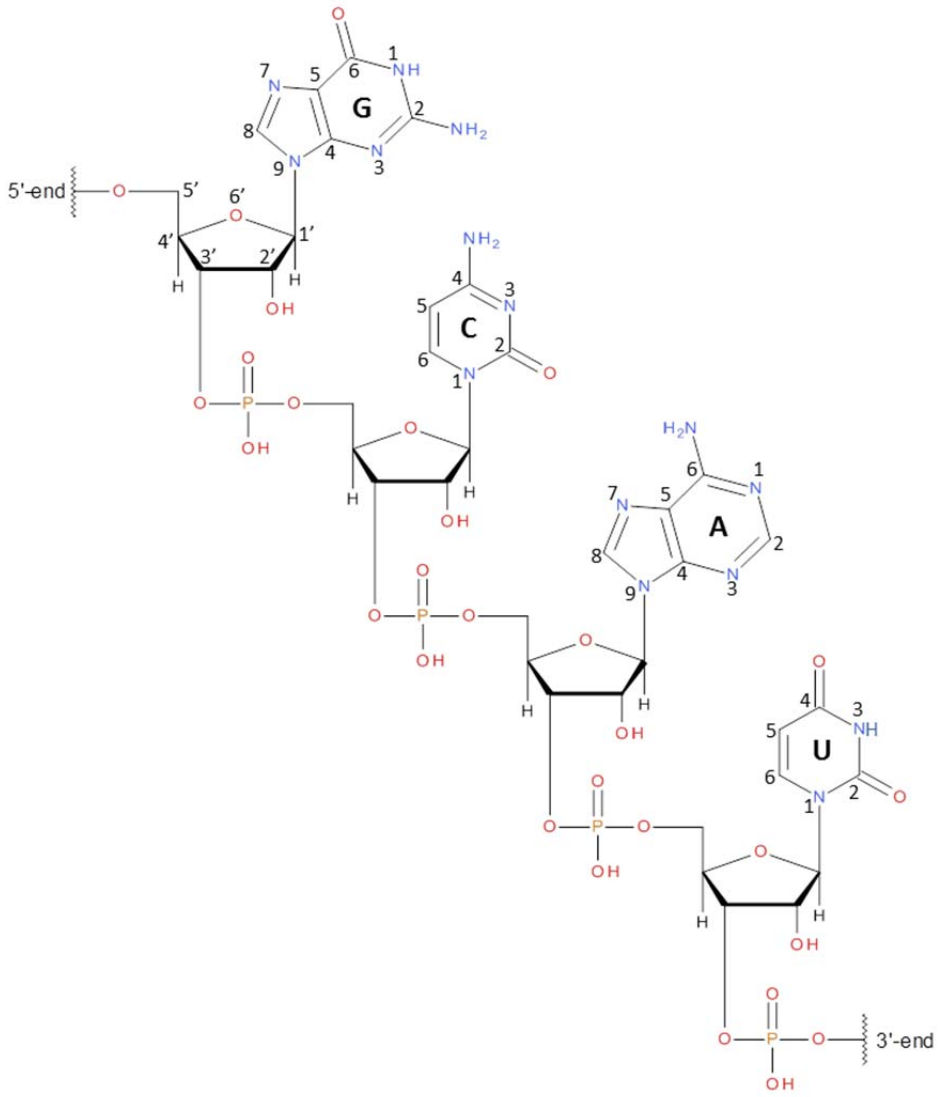


Figure 1: The primary structure of an RNA molecule. The primary structure of an RNA molecule is shown with four nucleotides - A, G, C and U. Also shown is the numbering system used to number each ringed atom within each base as well as in the ribose.

which it is a part of [30]; and to assemble the viral coat which encapsidates the hairpin loop-containing viral RNA [31]. Another example of a secondary structure element is the pseudoknot [32]. The pseudoknot is a complex structure comprising of two loops and two stems. Basically, one strand from the stem of a loop extends further and folds back to form intramolecular base pairs with nucleobases in the loop. This process creates a second loop and stem. Furthermore, these stems can stack on top of each other forming a quasi-continuous helix. In general, pseudoknot is a name given to a structurally diverse group that can vary in the length and number of loops and stems it forms. Functional roles of pseudoknots include the formation of

the catalytic core of ribozymes [33] and of self-splicing introns [34]. The pseudoknot is also a structural component of viral mRNA and is involved in ribosomal frameshifting during translation [35].

In this study, we will focus on the kink-turn motif – a secondary structure element found in double-stranded RNA [36]. Indeed the double-stranded RNA of which the kink-turn is a part of can be formed by the hybridisation of two individual RNA strands or alternatively by the self-hybridisation of a folded single RNA strand.

2.3 The Kink-Turn Structural Motif

The kink-turn, also called k-turn, is a RNA structural motif found in double-stranded RNA [36] which produces a tight kink in the axis of the RNA molecule. It is a widespread motif found in many functional RNA molecules [37]. It was first identified in rRNA when the structure of the 50S ribosomal subunit was solved [38]. Thereafter, it was found in many of the riboswitches, including SAM-I [39], T-box [40], cyclic-diGMP [41] and cobalamine [42], which adopt structures that can bind ligands with great specificity to control gene expression. Other than that, the k-turn motif is also found in small nucleolar RNAs [43-46], e.g. box C/D that forms a RNA-protein complex which functions to methylate rRNA [21]; and in small nuclear RNAs, e.g. in the spliceosomal U4 snRNA which is involved in spliceosome assembly [47, 48]. The k-turn motif is also found in the untranslated regions of mRNA [49], including the above mentioned riboswitches, which regulate the expression of the adjacent gene. Hence, the k-turn motif is a very common structural part of various RNA molecules involved in many cellular functions including ribosome assembly and structure, RNA translation, gene regulation, splicing and RNA modification. Additionally, putative k-turn forming sequences are also found in other functional RNA species [37] such as in ribonuclease P which is a ribozyme involved in tRNA processing [50]. Thus, the diversity of k-turn containing functional RNA molecules is ever increasing deeming the k-turn to be a critical secondary structure element in RNA.

2.4 Classification of Kink-Turns

K-turns can be divided into different classes depending on their sequence and structure. A simple k-turn is contained within a duplex RNA that has a central bulge, also called loop, generally comprising three unpaired nucleotides on the bulged strand followed by tandem G:A and A:G base pairs. The nomenclature used to label the nucleotide positions in the k-turn sequence [51] is shown in Figure 2. On folding the k-turn generates a tight kink in the axis of the RNA molecule.

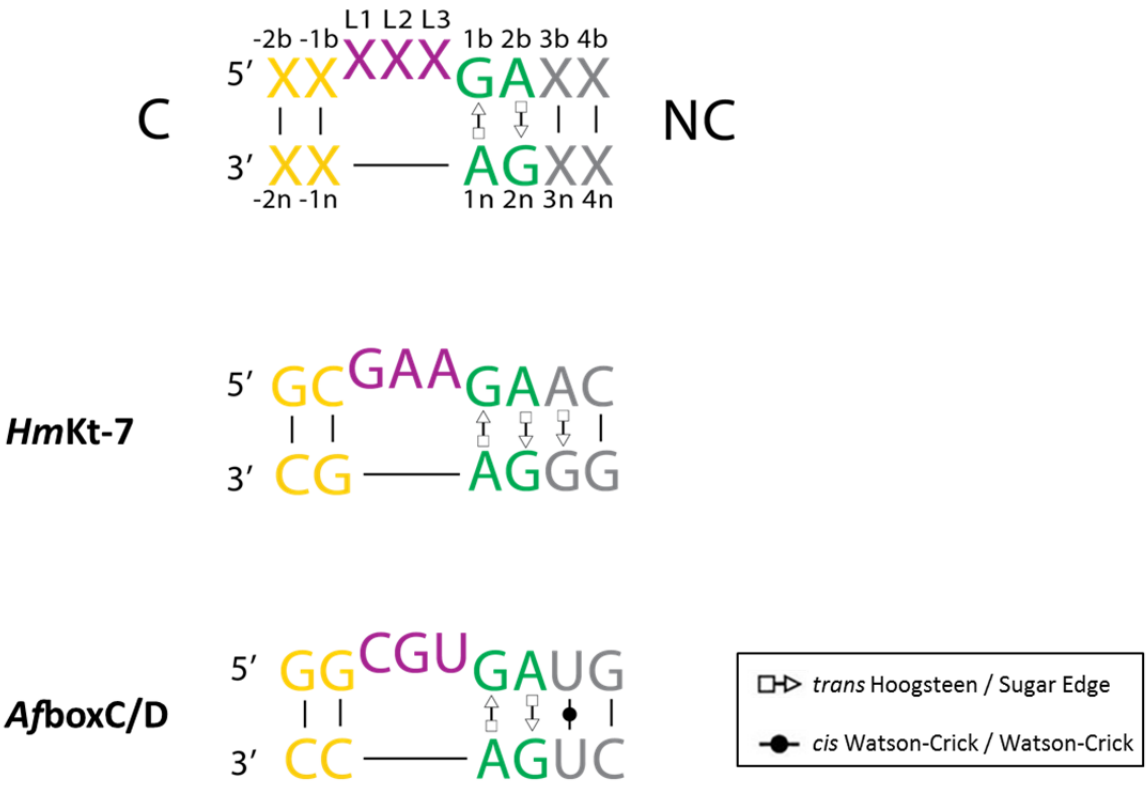


Figure 2: The nomenclature and sequence of standard kink-turns. The generic sequence of a standard simple k-turn with the numbering system is shown at the top. The G:A base pairs are highlighted green while other base pairs in the non-canonical (NC) helix are grey. The loop is highlighted magenta and the canonical (C) helix yellow. Please note that this colouring scheme is used throughout this work. The loop nucleotides are designated L*n* where *n* increases in number from 5' to 3'. The remaining nucleotides are numbered outwards from the bulge, positive for the NC helix and negative for the C helix. Nucleotides in the bulged strand take the suffix b while those in the non-bulged strand take n. The sequences of the standard k-turns *Hm*Kt-7 and *Af*boxC/D are shown.

The simple k-turn can be divided into standard and non-standard classes (Fig. 3) [36, 37]. The simple standard k-turn has G:A and A:G base pairs in the 1b:1n and 2b:2n positions, respectively. The *Hm*Kt-7 rRNA and the *Af*boxC/D snoRNA are examples of simple standard k-turns (Fig. 2). The simple non-standard k-turn differs by the means of having a substitution in one of the G:A pairs. For example, rRNA Kt-23 from both *T. thermophilus* and *E. coli* has A:U in the 2b:2n position although it forms a normal k-turn structure [52].

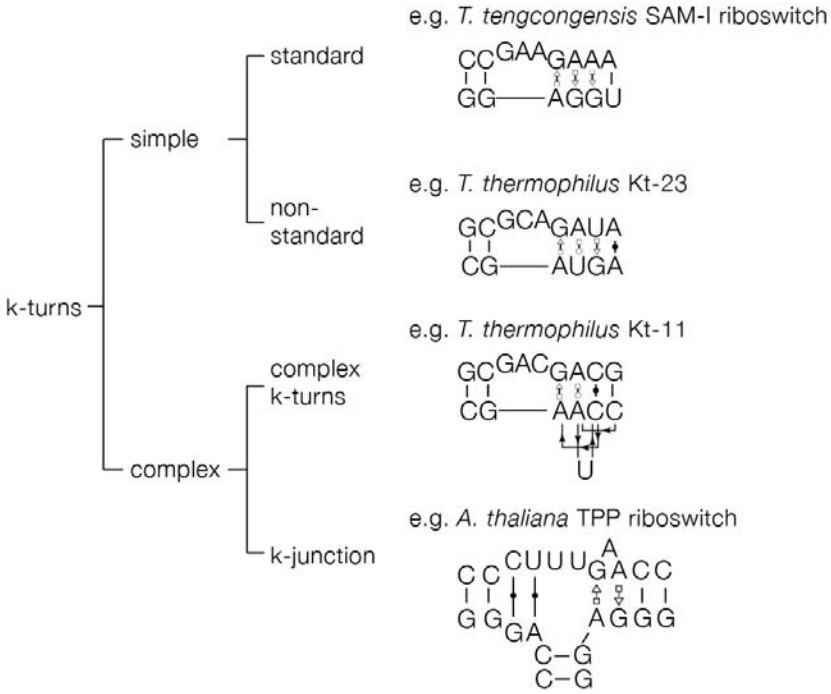


Figure 3: Classification of kink-turns. K-turns can be divided into simple and complex classes. The simple k-turns can be further divided into standard and non-standard classes with either the G:A and A:G base pairs in the 1b:1n and 2b:2n positions conserved or substituted, respectively. The primary sequence of complex k-turns does not map linearly on to the secondary structure. K-junctions, which are also complex, have a third helix inserted at the junction of the C and NC helices on the non-bulged strand. An example of each k-turn type is shown [37].

Other k-turns are regarded as complex (Fig. 3) when the nucleotides contributing towards the 1b:1n and 2b:2n positions cannot be mapped linearly onto the primary sequence. Yet the critical nucleotides, especially the conserved A1n and A2b are located at their normal positions within the structure, as discussed below. For example, in the primary sequence of *T. thermophilus* Kt-11, the 1n and 2n positions are separated by two

nucleotides one of which is C3n. In this case, the non-bulged strand of the NC helix doubles back on itself to form an S-turn. Nevertheless, A2b is placed normally in the k-turn structure where it hydrogen bonds to the -1n position (see next section) stabilising the kinked conformation. Another class of complex k-turns has also been identified called the k-junction [53]. In this case, the non-bulged strand is intercepted by a third helix to form a three-way helical junction. Nevertheless, the k-turn is normally developed. The different types of k-turns are summarised in Figure 3. In this study, we will focus on simple standard k-turns.

2.5 The Structure of a Simple Standard Kink-Turn

The k-turn produces a kink in the axis of the RNA molecule with an included angle of about 50˚ between the non-canonical (NC) and canonical (C) helices (Fig. 4) [36]. The NC helix is positioned to the 3' side of the loop and comprises the two critical G:A and A:G base pairs in the 1b:1n and 2b:2n positions, respectively. The C helix is positioned to the 5' side of the loop and, unlike the NC helix, consists of regular Watson-Crick base pairing throughout.

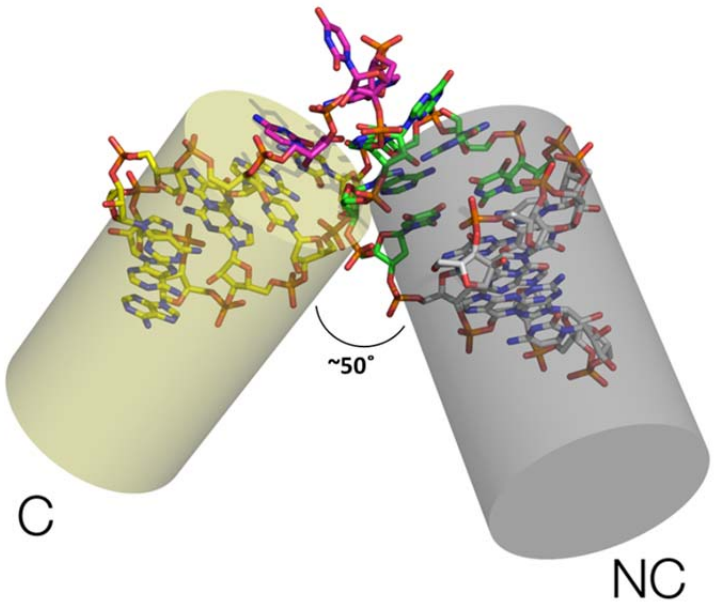


Figure 4: The overall structure of a standard kink-turn. The structure of a folded k-turn is displayed showing the strong kink in the axis of the RNA molecule with an angle of 50˚ [37].

In the NC helix both of the G:A base pairs in positions 1 and 2 are *trans* G(sugar edge):A(Hoogsteen edge) base pairs (Fig. 5) which are normally connected by two hydrogen bonds - GN2 to AN7 and AN6 to GN3 [37, 54]. The 2b:2n base pair is more variable in the number of hydrogen bonds it adopts omitting the AN6 to GN3 hydrogen bond in the N1 class k-turn structures (see next section). Nevertheless, the two G:A pairs and their hydrogen bonds are critically important for the stabilisation of the folded k-turn structure such that substitution of the guanine base in either position to an inosine, which prevents the GN2 to AN7 hydrogen bond from being formed, inhibits the metal ion-induced folding of the kink-turn [54].

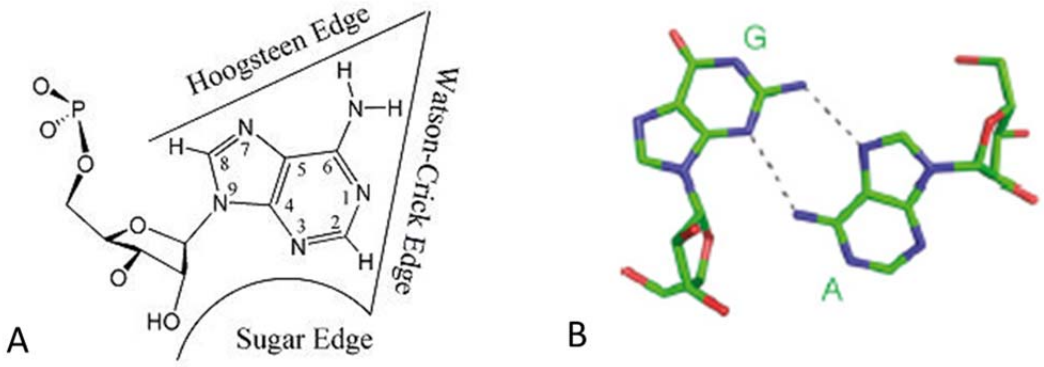


Figure 5: The *trans* sugar edge-Hoogsteen edge G:A base pair. A. The Hoogsteen, Watson-Crick and sugar edges of a nucleotide (in this case adenine) are shown. B. The *trans* G:A sugar edge (G) to Hoogsteen edge (A) base pair. The dashed lines show the two hydrogen bonds normally formed by the base pair from the GN2 to AN7 and AN6 to GN3 [55].

The k-turn conformation is stabilised by a number of interactions (Fig. 6). Firstly, the 5' nucleotide of the loop (L1) stacks on to the end of the C helix, while L2 stacks on to the end of the NC helix [37]. These stacking interactions stabilise the k-turn structure such that removal of the L1 and L2 nucleobases prevents the metal ion-induced folding into the kinked conformation (Appendix Fig. 1A, 1B). L3, on the other hand, protrudes away from the k-turn structure and does not take part in any such stabilising interaction. Thus, the removal of the L3 nucleobase does not impair the metal ion-induced folding of the k-turn (Appendix Fig. 1C). Upon folding the

minor grooves of the C and NC helices are juxtaposed, with the sugar edges of A1n and A2b directed towards the C helix. Key cross-strand hydrogen bonding interactions occur at the interface between the C and NC helices [51, 54, 56]. A-minor interactions [57] form between the adenine nucleobases and 2'-hydroxyl groups in the backbone of the C helix. The O2' of the ribose of L1 donates a hydrogen bond to A1nN1 [37]. The importance of this hydrogen bond has been investigated by the removal of the O2' by the substitution of L1 with the deoxyribose version of the nucleotide [51]. This completely prevented metal ion-induced folding of *Hm*Kt-7. A second critical hydrogen bond is donated by the -1nO2' to one of the ringed nitrogen atoms of A2b [37]. The acceptor can be either N1 or N3, creating two structural classes into which all k-turns can be divided. A third hydrogen bond observed in many k-turns is donated by the O2' of the ribose of L3 to the *pro*S non-bridging O of the phosphate linking L1 and L2 [37]. Substitution of L3 with the deoxyribose version of the nucleotide in *Hm*Kt-7 resulted in impaired metal ion-induced folding of the k-turn [51]. Moreover, additional sequence-dependant hydrogen bonds form adventitiously in particular k-turns.

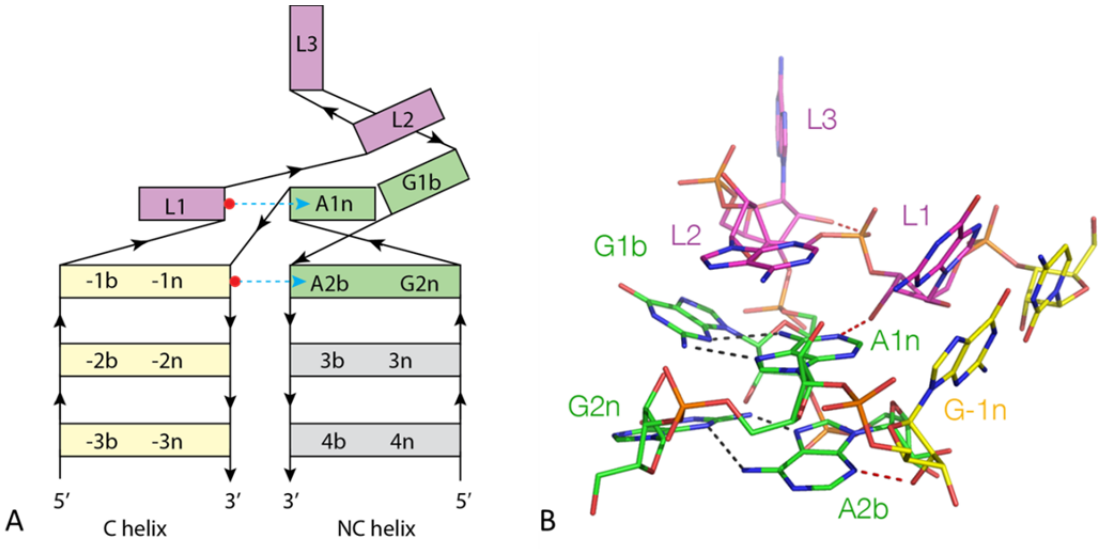


Figure 6: A-minor interactions in a folded standard kink-turn. A. The schematic of a standard folded k-turn is shown. The strong kink in the axis of the RNA molecule at the loop is displayed which directs the A nucleobases of the G:A and A:G base pairs in the 1b:1n and 2b:2n positions towards the C helix where they form critical cross-strand hydrogen bonds (dashed arrows) with the O2' of L1 and -1n nucleotides, respectively. The L1 and L2

nucleotides stack on to the ends of the C and NC helices, respectively [37]. B. The structure of the folded *Hm*Kt-7 is shown with hydrogen bonds as dashed lines. The two key cross-strand hydrogen bonds are donated by the O2' of L1 and -1n to the A1nN1 and A2bN3, respectively. These bonds, along with the bond from O2' of L3 to the *pro*S O of L1/L2 phosphate, are highlighted in red [37].

2.6 The N1 and N3 Classes of Kink-Turns

The ability of -1n O2' to donate a hydrogen bond to either the N1 or N3 position of A2b has resulted in the k-turns being split into these two respective conformational classes (Fig.7). Examples of N3 class k-turns include the SAM-I [39] and cyclic-diGMP [58] riboswitches, U4 snRNA [47], *Hm*Kt-46 [59] and box C/D snoRNA [44]. N1 class k-turns include the cobalamine riboswitch [42] and also *Hm*Kt-38 and *Hm*Kt-7 [59]. In these two alternative structural conformations, the rotational setting of the A2b nucleobase is altered which in turn affects its base pairing with G2n. While the GN2 to AN7 hydrogen bond is observed in both structural classes for the 2b:2n base pair, the AN6 to GN3 distance is typically >4Å in k-turns of the N1 class [37], i.e. there is no hydrogen bond. The N3 or N1 conformation formed is dependent on the local sequence. The base pair in the 3b:3n position has been shown to determine which class a k-turn shall adopt [60]. Nevertheless, the environment of the k-turn may also play a role, as the same sequence in different environments can adopt different conformations [60]. The major example of this is the simple standard k-turn, *Hm*Kt-7, which adopts an N1 class structure in the ribosome [59]. However, when placed in SAM-I riboswitch [56] or studied as a simple duplex RNA with or without the bound L7Ae protein [61], it adopts an N3 class structure. It is possible that the k-turn exists in a dynamic equilibrium between the two classes in free solution and it is only in larger complexes like a riboswitch that tertiary interactions fix it into a certain class.

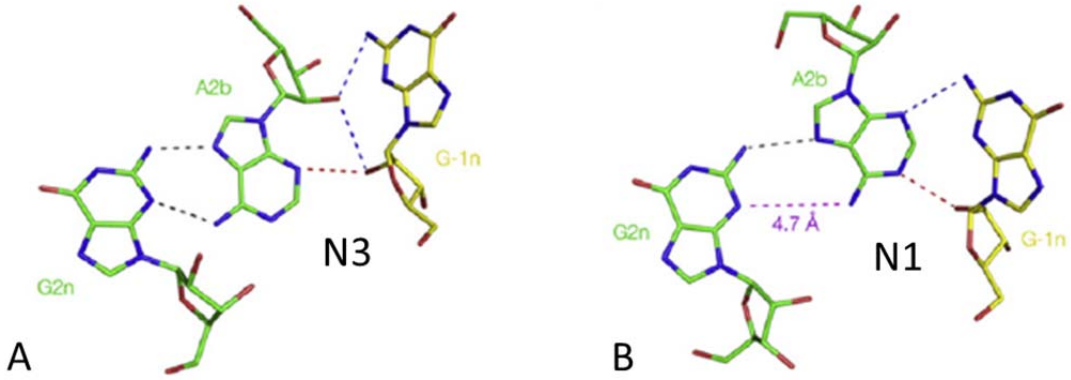


Figure 7: The A-minor interaction between the -1n and 2b positions in the N3 and N1 structural classes. The interaction between -1n and the 2b:2n base pair in a typical N3 (A) and N1 (B) structure is shown. The critical cross-strand hydrogen bond donated by O2' of G-1n to N3 (A) or N1 (B) of A2b is shown as a red dashed line. Other hydrogen bonds formed between -1n and 2b as well as within the 2b:2n base pair are shown as blue dashed lines. In the N1 structure (B), the distance between A2bN6 and G2nN3 is 4.7Å (shown as magenta dashed line) which is too long to form a hydrogen bond [36].

We have seen how hydrogen bonding differs between the N1 and N3 classes including the absence of the AN6 to GN3 hydrogen bond in the 2b:2n base pair in the N1 class. In *Hm*Kt-7, however, another difference in hydrogen bonding has been observed between the two conformational classes. In the N1 class, the 2'Os of the ribose rings in the -2n and 3b positions are hydrogen bonded [37, 51]. However, this interaction is absent from the N3 class as the distance between the two positions is far too great to allow a potential hydrogen bond to be formed. This suggests that in the two classes, the C and NC helices are positioned differently with respect to each other, i.e. the trajectory of the helices is altered when shifting from one structural conformation to the other. Undoubtedly, this conformational difference between the N3 and N1 structures will have implications on the tertiary interactions that are made by the helices of the k-turn [56]. Conversely, the tertiary interactions may influence the structural class that the k-turn adopts. This is exemplified by the example of *Hm*Kt-7 which, as described above, prefers to form an N3 structure when as a simple duplex RNA in free solution, but when in the ribosome it is forced to adopt the N1 structure. In the ribosome, the terminal loop of the helix which contains Kt-7 takes part in a tertiary interaction with a receptor, and the k-turn itself is also

bound to a protein called L24 [37]. It is possible that either or both of these interactions switch the conformation of Kt-7 from the N3 to the N1 class when in the ribosome.

2.7 The Folding of Kink-Turns

In free solution a k-turn sequence-containing RNA duplex exists in a conformational equilibrium between the folded k-turn state and an extended structure that resembles a normal three nucleotide bulge in a simple RNA molecule [62]. In the absence of metal ions and binding proteins, the equilibrium is biased towards the unfolded state [62, 63]. The equilibrium can be shifted towards the folded state, i.e. the population of the folded k-turn species can be increased, in three ways – by the addition of metal ions, protein or by tertiary interactions.

2.7.1 Metal Ion-Induced Folding of Kink-Turns

The addition of metal ions can induce the folding of k-turns [51, 62]. This can be studied by attaching fluorophores to the 5'-termini of the bulged and non-bulged strands of the k-turn containing duplex RNA of ~30 base pairs. Fluorescence resonance energy transfer (FRET) occurs between the fluorophores from the donor to the acceptor which is measured. The efficiency of FRET (E_{FRET}) depends on the inverse sixth power of the distance (R) between the fluorophores according to the Förster equation [64]:

$$E_{FRET} = [1+(R/R_0)^6]^{-1}$$

where R_0 is the Förster length, which is 56Å for the fluorescein donor-Cy3 acceptor pair [65] that is used in this work. Therefore, as the k-turn folds, the end-to-end distance between the two 5'-terminally attached fluorophores shortens resulting in an increase in the E_{FRET} [62]. The metal ion-induced folding of the k-turn can be followed by observing the increase in the E_{FRET}, which is measured in the steady-state, as a function of metal ion concentration. The resulting data are fitted to a two-state process allowing

the calculation of the apparent binding affinity, cooperativity and the midpoint ionic concentration at which folding is half-complete ($[M^{n+}]_{1/2}$). *Hm*Kt-7 is an example of a k-turn that folds well in metal ions exhibited by a large increase in E_{FRET} [51]. Both divalent, e.g. Mg^{2+}, and monovalent, e.g. Na^{+}, ions can be used to induce k-turn folding but a much higher concentration of the latter is required. For example, half-complete folding of Kt-7 occurs with $[Mg^{2+}]_{1/2}$ = 78μM or $[Na^{+}]_{1/2}$ = 30mM [51]. In this study Mg^{2+} ions are used to observe metal-ion induced folding of k-turns.

The reason for metal ions to be able to induce k-turn folding lies in their ability to bind to the k-turn structure stabilising it in the process. For example, crystallisation study of *Hm*Kt-7 as a free duplex has revealed a k-turn structure with two metal ions bound [66]. While one metal ion is coordinated to the O6 atoms of both G2n and G3n, the second metal ion is coordinated to only the G3nO6 atom (Fig. 8). Thus, metal ions stabilise the kinked conformation of *Hm*Kt-7 by coordinating to the O6 atoms of G2n and G3n in the major groove of the NC helix. Hence, in the case of *Hm*Kt-7, the G2n and G3n positions, and in particular their O6 atoms, can be deemed critical for metal ion-induced folding. This has been further confirmed by comparative gel electrophoresis [67] wherein the removal of the O6 atom from either of the two positions, G2n or G3n, in *Hm*Kt-7 resulted in less retarded electrophoretic mobility of the modified k-turn compared to the unmodified k-turn suggesting the impairment of k-turn folding despite the presence of Mg^{2+} ions [66]. Thus, overall, both 2b:2n and 3b:3n positions can be regarded as being important for metal ion-induced folding of k-turns. However, the 2b:2n position is conserved in all standard k-turns unlike the 3b:3n position [37]. Thus, it is the 3b:3n position that is the critical determinant for the folding of standard k-turns in metal ions.

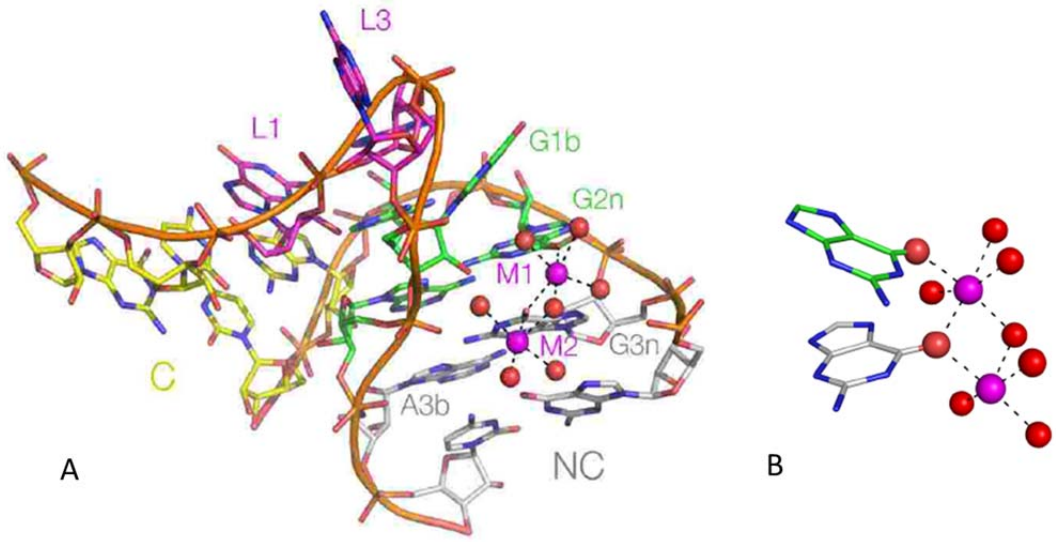


Figure 8: Mg^{2+} ions bound to the 2n and 3n positions of the kink-turn. A. The overall structure of the k-turn core with two Mg^{2+} ions, M1 and M2 (magenta spheres), bound to the G2n (green) and G3n (grey) positions is shown. B. A closer view reveals M1 coordinated to the O6 atoms of both G2n and G3n, while M2 is shown to coordinate only to the O6 atom of G3n [66].

2.7.1.1 The 3b:3n Sequence Determinant for Kink-Turn Folding in Metal Ions

Base pairs in the 3b:3n position that place a G into the 3n position shall be preferred for metal ion-induced folding of k-turns due to the ability of G3n to coordinate metal ions to stabilise the k-turn structure. *Hm*Kt-7, as discussed before, is a good example of this which has A3b:G3n [51, 66]. This proposition has further been confirmed by FRET experiments of 3b:3n variants of *Hm*Kt-7. Systematic modification of the 3b:3n position with all the other fifteen possible base pairs has revealed the preference of base pairs that place a G in the 3n position or, surprisingly, a C in the 3b position [66]. These allow good folding of the k-turn comparable to the unmodified *Hm*Kt-7. The exception to this is the normal Watson-Crick C:G base pair in the 3b:3n position which prevents folding suggesting the importance of non- Watson-Crick base pairing in this position. Two other base pairs, A:C and U:U, also allow moderate folding to occur of the k-turn. All other base pairs in the 3b:3n position, inhibit the metal ion-induced folding of the k-turn. Thus, in general,

the 3b:3n grid devised by McPhee *et al* (Appendix Fig. 2) can be used to predict the folding of standard k-turns in metal ions.

To test the theory implicated by the 3b:3n grid, the human U4 snRNA k-turn had been analysed by FRET [66]. The wild-type sequence which has G3b:C3n showed no increase in E_{FRET} upon the addition of Mg^{2+} ions, i.e. was unable to fold. However, the substitution of the 3b:3n position to an A:G base pair resulted in the Mg^{2+} ion-induced folding of the k-turn comparable to *Hm*Kt-7. Hence, the human U4 snRNA k-turn conforms to the 3b:3n rule. However, one example of a standard k-turn containing RNA molecule that does not conform to the 3b:3n rule is the *Af*boxC/D snoRNA. The *Af*boxC/D k-turn sequence has a U:U base pair in the 3b:3n position which, according to the 3b:3n grid, should allow moderate metal ion-induced folding to occur. However, this is not the case. No increase in E_{FRET} had been observed upon the addition of Mg^{2+} ions to the *Af*boxC/D k-turn RNA [66] (also this work). This suggested that, besides 3b:3n, there are other sequence determinants for the metal ion-induced folding of standard k-turn RNA which are undiscovered, leading to this work.

2.7.2 Kink-Turn Folding by Tertiary Interactions

Many k-turns are contained within large RNA molecules that fold into complex structures such as the ribosome. Within these structures, the k-turn motif mediates tertiary interactions with other parts of the structure, stabilising the overall structure. Conversely, these tertiary interactions also stabilise the k-turn in its folded conformation [37, 68].

The SAM-I riboswitch is an example of a simple standard k-turn containing large RNA molecule wherein the folded k-turn mediates tertiary interactions that stabilise the overall structure of the riboswitch, and, in the process, create a S-adenosyl methionine (SAM) ligand binding pocket [39, 69]. Consequently, the SAM-I riboswitch has been used a system to study k-turn RNA folding by the indirect means of ligand binding due to the direct connection between folding and binding. Only when the k-turn structure is

properly folded will the riboswitch adopt the correct structure, and only then will it bind its ligand. The enthalpy of ligand binding results in heat evolution that can be detected by isothermal titration calorimetry (ITC) [37].

For example, the A1nC substitution in the k-turn sequence of the SAM-I riboswitch resulted in no ligand binding [68]. This suggested the impairment of k-turn folding and subsequently of the riboswitch such that it cannot bind its ligand. On the other hand, a G2nA substitution in the k-turn sequence of the SAM-I riboswitch allowed normal ligand binding to occur. Crystallisation studies have also revealed the G2nA modified SAM-I riboswitch to adopt a similar structure to the unmodified riboswitch [68], i.e. the k-turn and consequently the riboswitch structure fold normally despite the G2nA modification. However, the same G2nA modification in the k-turn RNA as a simple duplex prevented metal ion-induced folding [68]. Thus, the same k-turn RNA sequence behaves differently in the two environments. Evidently the tertiary contacts in the riboswitch provide additional free energy that stabilises the modified k-turn which otherwise is unable to fold in metal ions alone.

2.7.3 Kink-Turn Folding by Protein Binding

Many kink-turns act as binding sites for proteins, and in principle the free energy of binding could be coupled to the RNA folding. This was demonstrated for the binding of *A. fulgidus* L7Ae to *Hm*Kt-7 [54]. L7Ae is a widespread family of proteins that bind to k-turns in the ribosome, in snoRNP, in the spliceosome and even in a ribozyme. *Af*L7Ae binds to *Hm*Kt-7 with pM affinity, resulting in folding of the k-turn [70].

The L7Ae family of proteins are ubiquitous and are found in all types of organisms. Examples in eukaryotes and archaea include the proteins L7Ae, L30e and S12e [71], as well as the human 15.5 kDa (also called 15.5k) protein [72]. In yeast, examples include the Nhp2 and Snu13p proteins [37, 73]. Bacteria also have analogs, such as the YbxF protein [74]. Each of

these proteins is capable of binding k-turn containing RNA molecules although they can be selective in which k-turn sequences they bind. For example, the binding of *Af*L7Ae will result in the folding of many k-turns, such as the *Hm*Kt-7 and *Af*boxC/D k-turns in this work. However, the *Af*L7Ae protein is unable to bind to the human Kt-23 and Kt-42 k-turns, leading to no increase in E_{FRET} (Appendix Fig. 3). Interestingly, the difference between these two sets of sequences lies in the L2 position which is involved in the protein binding interaction (discussed later). *Hm*Kt-7 and *Af*boxC/D both have a purine base in the L2 position, the identity of which is A and G, respectively. The human Kt-23 and Kt-42 k-turns, on the other hand, have the pyrimidine U in the L2 position of the k-turn sequence. This reveals the k-turn sequence selectivity of the L7Ae protein. And, overall, this provides a plausible explanation for the existence of a family of related proteins which primarily function as k-turn binding proteins but differ with respect to the k-turn sequence they bind.

Nevertheless, the RNA-protein complex formed upon binding of L7Ae, or a related protein, to a k-turn containing RNA molecule, can have various downstream functions. For example, the box C/D and H/ACA snoRNP complexes, the assemblies of which are initiated by the binding of an L7Ae-type protein to the k-turn motifs flanking the guide RNA, are involved in the site-specific 2'-*O*-methylation and pseudouridylation of rRNA, respectively [21]. The U3 snoRNP complex also recruits the L7Ae protein and is involved in the cleavage of pre-rRNA into the mature rRNA [73]. Additionally, the U4-U6·U5 tri-snRNP complex binds the 15.5k protein to assemble the spliceosome [47, 72]. L7Ae is also a component of archaeal RNaseP which is involved in tRNA maturation [75]. Thus, the L7Ae family of proteins play a number of important roles as a complex with the k-turn containing RNA molecule, from the assembly of ribosomes and spliceosomes, to the site-specific modification of RNA.

2.7.3.1 Binding Interaction between the L7Ae Protein and Kink-Turn RNA

Various structures of protein bound to k-turn RNA have been determined including *H. marismortui* L7Ae-Kt-15 [59]; *A. fulgidus* L7Ae-boxC/D k-turn [44]; human 15.5k protein-U4 snRNA k-turn [47]; and *Af*L7Ae-*Hm*Kt-7 [61]. The comparison of all these structures has revealed some common protein binding interactions which form the basis of the recognition of k-turn structures by the L7Ae proteins [37].

Proteins bind to the outside of the structure in the major groove of the k-turn (Fig. 9A). Two sections of the proteins form the binding interface. These are an α-helix element and a short loop of hydrophobic residues. The α-helix enters the major groove of the NC RNA helix where its N-terminal end makes specific contacts with the major groove edges of the G1b and G2n nucleobases [37]. For example, in the *Af*L7Ae-*Hm*Kt-7 complex, glutamate 34 at the N-terminus of the α-helix is hydrogen bonded to G1bN1 in the k-turn RNA (Fig. 9B). Additionally, the adjacent asparagine 33 is hydrogen bonded to G2nO6. Furthermore, G1bO6, which is partially negatively charged, is positioned close to the axis of the α-helix at the N-terminus such that it is electrostatically stabilised by the positive pole of the helix dipole, which is another universal interaction. The C-terminus of the α-helix is situated close to the non-bulged strand of the NC RNA helix and takes part in non-specific interactions [37]. Thus, it is the α-helix N-terminus that is important in the sequence-specific recognition of the core of the k-turn. The hydrophobic loop caps the L1 and L2 positions of the k-turn RNA making good Van der Waals interactions that bury 732Å^2 [37]. However, it is the L2 position that makes the maximal contact. In the *Af*L7Ae-*Hm*Kt-7 complex, isoleucine 88 is situated directly over the L2 position (Fig. 9C), with a similar interaction found in other complexes. Also, glutamate 89 within the loop makes a specific hydrogen bond to GL1 N1. Altogether, these binding interactions allow the specific molecular recognition, and subsequent binding, of the k-turn structures by the L7Ae proteins.

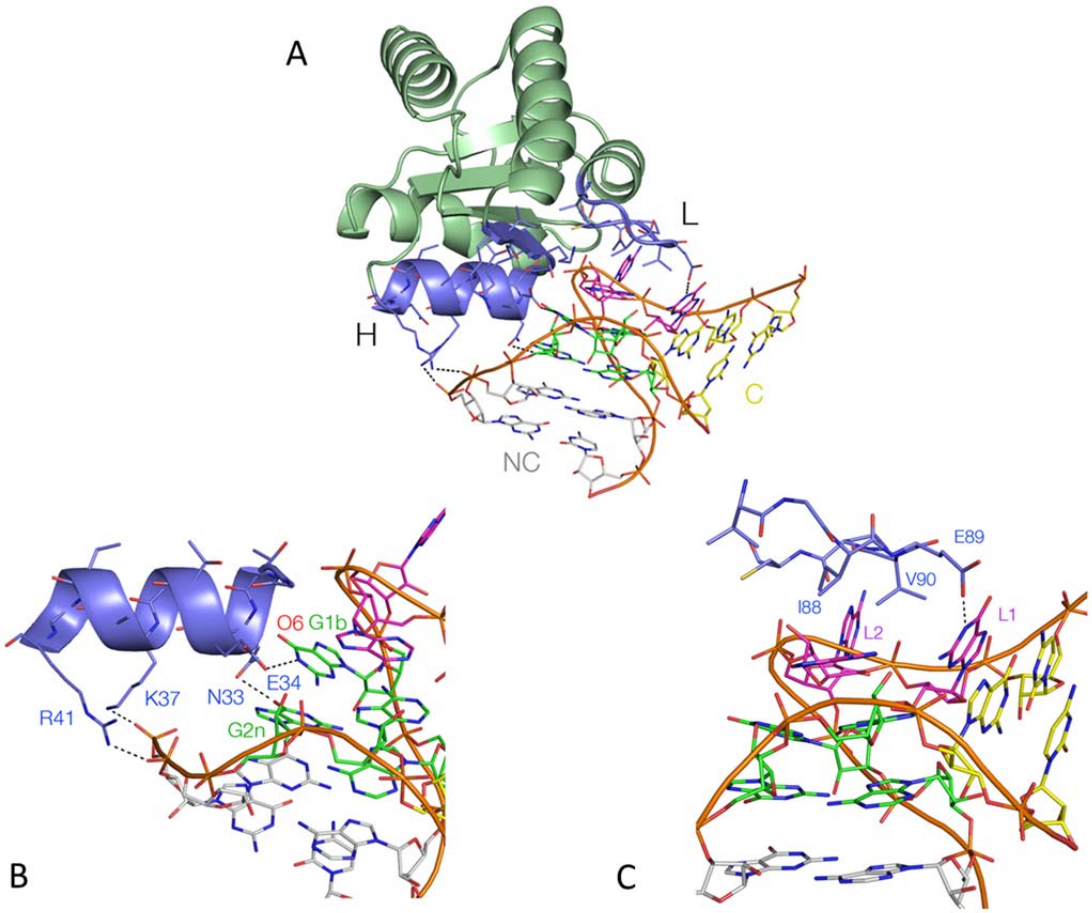


Figure 9: Structure of a complex between *H. marismortui* Kt-7 and *A. fulgidus* L7Ae. A. The overall structure of L7Ae (shown in cartoon form) bound to the outer face of Kt-7 in the major groove. The two RNA binding elements of L7Ae, the α-helix (H) and the hydrophobic loop (L), are highlighted in blue. B. The interaction of the α-helix with the major groove of the NC helix is shown with hydrogen bonds as dashed lines. The glutamate (E34) and asparagine (N33) make specific hydrogen bonds to G1b N1 and G2n O6, respectively. The lysine (K37) and arginine (R41) make non-specific contacts with the backbone of the NC helix. The electronegative oxygen G1b O6 is located at the positive pole of the helix dipole of the α-helix. C. The hydrophobic loop caps the loop region of the k-turn, sitting over the L1 and L2 nucleobases making good Van der Waals contacts. L2 is in direct contact with isoleucine (I88). Glutamate (E89) hydrogen bonds to GL1 N1.[37]

2.7.3.2 Assembly of the Box C/D snoRNP Complex

Box C/D snoRNA is an example of a k-turn containing RNA molecule which folds into the kinked conformation upon binding of the L7Ae-type protein [44, 66] (also this work). This is the first step in the assembly of the biologically

active box C/D snoRNP complex which is involved in the site-specific 2'-*O*-methylation of rRNA [21, 76-81]. Box C/D snoRNA comprises of a bulged RNA duplex in which the central 12 nucleotides of each bulged strand provide the guide sequences which are complementary to the target RNA molecules that are to be methylated. Thus, the guide sequences hybridise to the target RNA molecules with specificity. Flanking the guide region at each end are duplexes that contain the box C/D and box C'/D' sequences. Both of these sequences contain the k-turn sequence motif and are able to adopt the k-turn conformation.

On this snoRNA molecule many proteins assemble to form the snoRNP complex [82-85]. In the first step an L7Ae-family protein (e.g. 15.5k in humans), binds to each of the two k-turn sequences stabilising their kinked conformation. This is followed by the binding of the Nop58 protein to box C/D and the Nop56 protein to box C'/D' in humans (Nop5 proteins in archaea). Lastly, two molecules of the SAM-dependent methyl transferase fibrillarin are recruited to generate the catalytically-active complex (Fig. 10). Thus, the key step is the binding of the L7Ae-type protein to the k-turn which initiates the assembly of the complex. It has been shown that modifications in the k-turn sequence, that prevent its folding into the kinked conformation, inhibit the assembly of the complex [83]. For example, mutations in the critically important G:A and A:G base pairs in the 1b:1n and 2b:2n positions of the k-turn sequence, respectively, prevent k-turn folding and therefore the binding of the 15.5k protein to the human box C/D RNA [83]. As a result, the assembly of the box C/D snoRNP complex is blocked. Interestingly, recent studies have revealed a naturally occurring modification in human box C/D k-turn RNA that impairs its folding into the kinked conformation, thus prevents the binding of the 15.5k protein, which is N^6-adenine methylation [55] (discussed later).

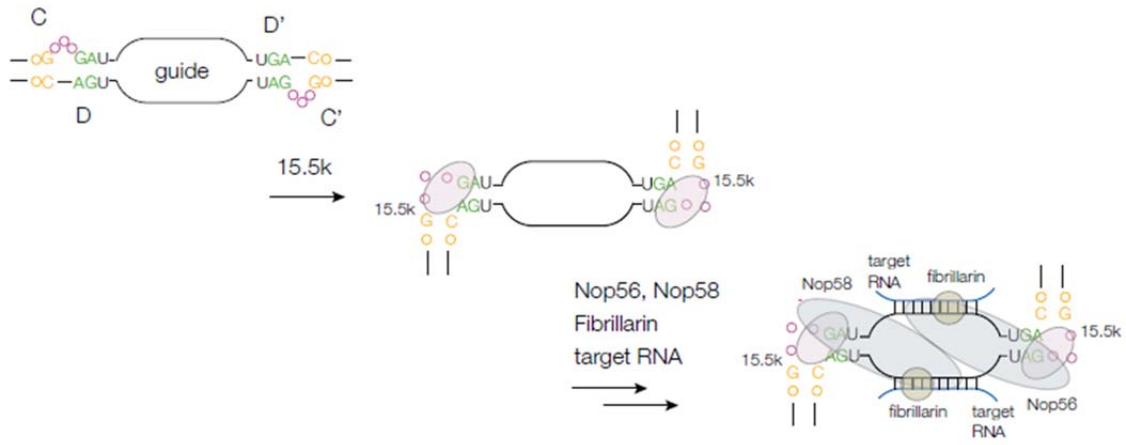


Figure 10: Assembly of the human box C/D snoRNP complex. The first step is the binding of the 15.5k protein to the k-turn sequence motifs in the box C/D and box C'/D' regions of the snoRNA, to generate the k-turns. Subsequently, the Nop58 and Nop56 proteins bind to the box C/D and box C'/D' regions, respectively, and two molecules of fibrillarin are recruited to generate the catalytically active box C/D snoRNP complex. The guide sequences bind to their complementary target RNA to undergo 2'-*O*-methylation.

2.8 RNA Modification

Dynamic covalent modification of cellular RNAs is diverse and widespread, being found in all forms of life [86]. More than 100 modification types are known to exist in the vast range of native cellular RNAs which include mRNAs, tRNAs, rRNAs, snRNAs, and also snoRNAs [55]. These post-transcriptional modifications of cellular RNAs are catalysed by specific RNA modification enzymes. Examples of such modifications include pseudouridylation and methylation. For example, the H/ACA snoRNP complex recruits a pseudo-U synthetase to carry out the site-specific pseudouridylation of the target rRNA [22, 23]. Similarly, the box C/D snoRNP complex recruits a methyl transferase to carry out the site-specific 2′-*O*-methylation of the target rRNA [21, 76]. As well as the ribose rings of RNA, the nucleobases are also modified post-transcriptionally [86]. We will focus on the methylation of the adenine base at the N^6 position – the N^6-methyladenine (m6A) modification (see next section).

2.8.1 The m6A Modification

N^6-methyladenine is the one of the most common modification in RNA [87-89]. It is found in coding RNA and is the most prevalent modification in eukaryotic mRNA. It is also found in non-coding RNA, such as tRNA [90] and rRNA [91]; as well as in other long non-coding (lnc) RNA, such as MALAT1 which is involved in the regulation of alternative splicing [11, 12]. Furthermore, viral RNA has also been found to contain this modification [92, 93]. The m6A modification is also found in snRNA [94] and, interestingly, has recently been found in snoRNA [55] (discussed later).

The N^6-methylation of adenine involves the transfer of a methyl group to the N^6 position of the adenine base by the methyltransferase enzyme [95] using SAM as the methyl donor [96]. The methyltransferase enzyme methylates the target RNA with specificity. It recognises the core sequence motif of GAC (70%) or AAC (30%), and catalyses the N^6-methylation of the central A nucleobase [97, 98]. Interestingly, the box C/D k-turn RNA sequence has the GAC core sequence motif making it a putative target for the m6A modification (see next section). The process of N^6-adenosine methylation is, however, reversible wherein the removal of the methyl group is catalysed by the demethylase enzyme [99, 100]. Together the methyltransferase and demethylase enzymes modulate the level of m6A modification in RNA. The deregulation of these enzymes, which leads to abnormal levels of the m6A modification, has been linked to diseases such as infertility [101], obesity [89, 102] and cancer [103].

The m6A modification in RNA has various downstream functions which are mediated via the regulation of protein binding. Firstly, the m6A-modified RNA can act as a site of recognition for methyl-specific binding proteins. For instance, the m6A-modified mRNA is recognised and bound by the human YTHDF2 protein which targets its degradation [97]. Thus, the m6A modification is involved in the regulation of mRNA stability. However, at the same time, the m6A modification can control the efficiency of mRNA translation [104]. In association with the human YTHDF1 protein, the m6A-

modified mRNA can interact with the translation machinery to promote the translation of the mRNA [105]. Overall, by coordinating the processes of mRNA degradation and translation, the m6A modification provides a way to regulate gene expression. Secondly, although the m6A modification mediates the downstream effects of mRNA degradation and translation promotion via protein binding, it can conversely prevent the interaction of a RNA binding protein with the m6A-modified RNA by perturbing the local RNA structure, thus affecting the function of the RNA concerned. The N^6-methylation of A disrupts its *trans* sugar edge(G)-Hoogsteen edge(A) base pairing with G [55], which is a critically important structural element of the k-turn [54] (see next section).

2.8.1.1 Disruption of the *trans* G(Sugar Edge): A(Hoogsteen Edge) Base Pair: Implications for the Box C/D Kink-Turn

As discussed previously, the k-turn containing box C/D snoRNA associates with proteins to form a snoRNP complex which is involved in the 2'-*O*-methylation of rRNA [81]. The first key step in the assembly of this complex is the binding of the L7Ae protein to the k-turn regions of the snoRNA to stabilise the k-turn conformation of the RNA which, otherwise, is unable to fold in the absence of protein in metal ions alone [44, 66, 83]. A process that inhibits the binding of the L7Ae protein to the k-turn RNA shall therefore prevent the assembly of the complex. Such a process has been identified which is the N^6-methylation of the adenine in the 1n position of the box C/D k-turn RNA [55].

The box C/D k-turn RNA can potentially have the GAC core sequence motif in the [2n, 1n,-1n] region which is specifically recognised and methylated by the methyltransferase enzyme. A1n and G2n positions are conserved in all standard k-turns [37] so the -1n position is the key to providing the critical GAC methylation sequence. Bioinformatic data mining has revealed a subset of human box C/D k-turn sequences that have a G:C base pair in the -1b:-1n position making them putative targets for the m6A modification at the 1n position [55]. And, indeed many of these box C/D k-turn sequences are

naturally found to be methylated at the 1n position with the m6A modification. Interestingly, A1n forms the critically important *trans* sugar edge(G)-Hoogsteen edge(A) base pair with G1b in the k-turn [54], and the m6A modification has been shown to disrupt this base pairing [55]. Hence, upon A1n N^6-methylation, the disruption of the *trans* sugar edge(G1b)-Hoogsteen edge(A1n) base pair leads to the box C/D k-turn RNA being unable to fold into the kinked conformation and bind the 15.5k protein as revealed by the FRET and gel electrophoresis experiments [55]. The inability to bind the 15.5k protein shall prevent the further assembly of the box C/D snoRNP complex [83]. Thus, by preventing the binding of the L7Ae-type protein to the box C/D k-turn RNA, the N^6-methylation of adenine provides a way to regulate the assembly of the functionally active box C/D snoRNP complex, which itself is involved in the methylation of RNA, revealing an interesting series of events.

3. Aims and Objectives

The 3b:3n base pair can determine the folding ability of k-turn RNA in metal ions. Based on the 3b:3n grid devised by McPhee *et al* (Appendix Fig. 2), it is predicted that k-turns that have the U3b:U3n sequence element can fold to a moderate degree in metal ions like the 3b:3n=U:U modified *Hm*Kt-7 k-turn. However, this is not the case as exemplified by the *Af*boxC/D k-turn which naturally has the U3b:U3n sequence element but fails to undergo metal ion-induced folding into the kinked conformation. This suggests the existence of other sequence elements in the k-turn which also influence its folding properties in metal ions. The comparison of the *Hm*Kt-7 and *Af*boxC/D k-turn sequences has revealed three other sequence elements that differ between the two k-turns, other than the 3b:3n base pair, which could potentially affect the folding properties of the k-turns. These are the -1b:-1n and 4b:4n base pairs and the loop. Therefore, in addition to the 3b:3n base pair, it is the effect of these other sequence elements on metal ion-induced k-turn folding that is studied in this work.

Folding properties in the presence of metal ions of the well folding *Hm*Kt-7 and unfolding *Af*boxC/D k-turns and their variants have been studied in this work. Folding is primarily studied by the technique of FRET, as well as comparative gel electrophoresis. The -1b:-1n, 3b:3n, 4b:4n and loop sequence elements are systematically exchanged between these two k-turns to determine their effect on k-turn folding in the presence of metal ions. Initially, one sequence element at a time is exchanged (single modifications) followed by double and triple modifications, until eventually all four sequence elements are exchanged converting the *Hm*Kt-7 sequence into the *Af*boxC/D sequence and vice versa. Expectedly, the systematic substitution of the *Hm*Kt-7 sequence elements with those from *Af*boxC/D will reduce its folding ability in metal ions, while replacement of the *Af*boxC/D sequence elements with those from *Hm*Kt-7 will increase its folding ability. Nevertheless, the magnitude of effect exerted by each sequence element and its relative importance in metal ion-induced folding of k-turns will be observable.

Furthermore, the effect of the N^6-methyladenine modification on k-turn folding is investigated. Adenine nucleobases in the 1n, 2b, 3b and L3 positions of the *Hm*Kt-7 k-turn sequence are studied. These positions in the *Hm*Kt-7 k-turn are individually modified to N^6-methyladenine and folding in the presence of metal ions and protein is observed by FRET, to determine the effect of each position modification on k-turn folding should these modifications be naturally found to occur in k-turn RNA.

Overall, in this work we seek to answer the following questions:

1. Is the metal ion-induced folding of k-turns in RNA dependant on their various sequence elements?
2. Does the N^6-methyladenine modification in the various positions of the k-turn affect its folding properties?

4. Materials and Methods

4.1 RNA Synthesis and Deprotection

The k-turn sequence-containing RNA oligonucleotides were chemically synthesised using the solid-phase phosphoramidite chemistry [106] implemented on the Applied Biosystems 394 DNA/RNA synthesiser machine. The 2'*O t*-BDMS-protected ribonucleotide phosphoramidites [107] used for the synthesis of the RNA are summarised in Table 1, along with all the reagents and consumables used for RNA synthesis and deprotection. The fluorophores, fluorescein and Cy3, were attached to the 5′ termini of the oligonucleotides as phosphoramidites in the final cycle of the synthesis, as required. Oligonucleotides for the FRET and comparative gel electrophoretic experiments were synthesised using the 0.2μmol and 1μmol synthesis columns, respectively. For comparative gel electrophoretic experiments, the DNA-RNA-DNA species were synthesised by incorporating deoxyribonucleotides (Table 1) into the oligonucleotide at the required positions. The sequences of the wild-type *H. marismortui* Kt-7 and *A. fulgidus* box C/D RNA synthesised for use in FRET experiments are detailed in Table 2, as well as the DNA-RNA-DNA sequence for the same unmodified *H. marismortui* Kt-7 synthesised for use in the comparative gel electrophoresis experiment. Modifications were introduced to these sequences as required.

The synthesised oligonucleotides were cleaved off column and base deprotected by the application of 2mL 25% ethanol/ammonia solution to the synthesis column for 3h at 20°C. Thereafter, the oligonucleotide samples were transferred to glass vials with screwed lids. The fluorophore-labelled oligonucleotides, as well as the DNA-RNA-DNA species, were further deprotected for 2h at 65°C. After deprotection, the oligonucleotides were cooled on ice for 10mins, transferred to 2mL eppendorf tubes and, with the lids opened or pierced, evaporated to dryness overnight by placement in a SpeedVac. They were re-dissolved in 115μL DMSO to which was added 60μL triethylamine (TEA) and 75μL 1M triethylamine trihydrofluoride (TEA, 3HF) and incubated at 65°C for 2.5h to remove the *t*-BDMS protecting groups. Thereafter, the samples were cooled on ice for 10 mins and 250μL

RNA Quenching Buffer was added. The oligonucleotides were then desalted by application to NAP-10 columns.

	Name	Manufacturer
RNA Phosphoramidites	Pac-A-CE Phosphoramidite (A)	Link Technologies Ltd.
	iPr-Pac-G-CE Phosphoramidite (G)	Link Technologies Ltd.
	Ac-C-CE Phosphoramidite (C)	Link Technologies Ltd.
	U-CE Phosphoramidite (U)	Link Technologies Ltd.
	I-CE Phosphoramidite (I)	Glen Research
	N6-Methyl-A-CE Phosphoramidite (m6A)	Glen Research
	rSpacer CE Phosphoramidite (S)	Berry & Associates
DNA Phosphoramidites	dA(tac) CE Phosphoramidite (A)	Proligo Reagents / Sigma-Aldrich
	dG(tac) CE Phosphoramidite (G)	Proligo Reagents / Sigma-Aldrich
	dC(tac) CE Phosphoramidite (C)	Proligo Reagents / Sigma-Aldrich
	dT CE Phosphoramidite (T)	Proligo Reagents / Sigma-Aldrich
Fluorophores	Cy3 Amidite	GE Healthcare Life Sciences
	5'-Fluorescein-CE Phosphoramidite	Link Technologies Ltd.
Synthesis Columns	0.2µmol dG^{dmf} 1000Å CPG DNA synthesis column	Applied Biosystems / Thermo Fisher Scientific
	1µmol dmf-dG Synbase™ CPG 1000/110S ALLFIT Column	Link Technologies Ltd.
Synthesis Reagents	ETT Activator (0.5M)	Link Technologies Ltd.
	Cap Mix A (Pyridine / Pac-Anhydride)	Link Technologies Ltd.
	Cap B	Sigma-Aldrich
	Oxidiser 0.02M	Sigma-Aldrich
	Acetonitrile	Sigma-Aldrich
	TCA Deblock	Sigma-Aldrich
	Dichloromethane	Applied Biosystems / Thermo Fisher Scientific
Deprotection Reagents / Consumables	Ammonium hydroxide solution	Sigma-Aldrich
	Absolute ethanol	VWR International
	Dimethyl sulfoxide (DMSO)	Sigma-Aldrich
	Triethylamine (TEA)	Sigma-Aldrich
	Triethylamine trihydrofluoride (TEA, 3HF)	Sigma-Aldrich
	RNA Quenching Buffer	Glen Research
	NAP-10 columns	GE Healthcare Life Sciences

Table 6: Phosphoramidites, reagents and consumables used in the synthesis and deprotection of oligonucleotides.

	Construct	Oligo Sequence (5' to 3')
FRET Experiment	Kt-7 wild-type bulged strand	Flu-CCAGUCAGUGGCGAAGAACCAUGUCAGG
	Kt-7 wild-type non-bulged strand	Cy3-CCUGACAUGGGGAGCCACUGACUGG
	box C/D wild-type bulged strand	Flu-CCUCAGUGGGCGUGAUGCAUGUCAUG
	box C/D wild-type non-bulged strand	Cy3-CAUGACAUGCUGACCCACUGAGG
Comparative Gel Electrophoretic Experiment	Kt-7 wild-type bulged strand	CGCAAGCGACAGGAACCTCGCCAGUCAGUGGCGA AGAACCAUGUCAGGGGACTGTCAAGTTGAACAGG
	Kt-7 wild-type non-bulged strand	CCTGTTCAACTTGACAGTCCCCUGACAUGGGGAGC CACUGACUGGCGAGGTTCCTGTCGCTTGCG

Table 7: Oligonucleotide sequences for wild-type *H. marismortui* Kt-7 and *A. fulgidus* box C/D synthesised for use in FRET and comparative gel electrophoresis experiments. The DNA sections of the oligonucleotides are underlined.

4.2 Oligonucleotide Purification

The oligonucleotides were purified by gel electrophoresis under denaturing conditions. The denature gel stock solution comprised of 20% acrylamide: bis-acrylamide (19:1) (Scientific Laboratory Supplies), as well as 7M urea, in 90mM Tris.borate (pH 8.5) and 10mM EDTA buffer (1X TBE). For pouring one gel, 500µL 10% APS and 25µL TEMED was added to 60mL of the denature gel stock solution. The gel was poured and allowed to set for at least 3h at RT. Thereafter, an equivalent volume of formamide was added to 80µg of the oligonucleotide sample, which was then heated at 95˚C for 2mins then cooled on ice for 5mins. The sample was then loaded into one lane of the gel. After loading the samples in the gel, electrophoresis was performed at 25W for ~3h with 1X TBE as the running buffer. Thereafter, nucleic acids were visualized by UV shadowing and bands corresponding to the full length products were excised, and electroeluted into 140µL 8M ammonium acetate in 1X TBE at 150V at RT using the EluTrap equipment. The nucleic acid was then precipitated with ethanol.

Fluorophore-labelled oligonucleotides were subjected to further purification by reversed-phase HPLC on a C18 column (ACE 10-300, Advanced Chromatography Technologies), using an acetonitrile gradient with an aqueous phase of 100mM triethylammonium acetate (pH 7.0) (Thermo Fisher Scientific). The samples collected were evaporated to dryness and re-suspended in 120µL ultrapure water.

4.3 RNA Hybridisation and Native Gel Purification (for FRET experiments)

Equimolar quantities of 200pmol of the appropriate oligonucleotides were annealed in 1X TBE and 25mM NaCl by slow cooling from 95°C to 4°C in a water bath. Thereafter, the hybridised RNA was kept cold. The hybridised RNA was purified by gel electrophoresis under non-denaturing conditions. The native gel solution comprised of 12% acrylamide: bis-acrylamide (29:1) (Scientific Laboratory Supplies), as well as 25mM NaCl, in 1X TBE. For pouring one gel, 583µL 10% APS and 29.2µL TEMED was added to 70mL of the native gel solution. Again, the gel was poured and allowed to set for at least 3h at RT. Thereafter, 2.5% ficoll was added to the hybrid RNA sample, which was then loaded into one lane of the gel. After loading the samples in the gel, electrophoresis was performed at 150V at 4°C for ~ 6 h, with 1X TBE and 25mM NaCl buffer recirculation. The gel was then visualised by UV shadowing and bands containing the duplex RNA were excised from the gel and electroeluted into 140µL 8M ammonium acetate in 0.25X TBE at 100V at 4°C, followed by ethanol precipitation and air-drying at 4°C.

4.4 Fluorescence Resonance Energy Transfer (FRET) Experiment

The FRET efficiency was measured for a series of RNA duplexes, the *H. marismortui* Kt-7 and *A. fulgidus* box C/D k-turn RNAs and their modified variants as indicated in the text. These were chemically synthesised and

terminally 5'-labelled with fluorescein donor on the bulged strand and Cy3 acceptor on the non-bulged strand. The sequences of the unmodified Kt-7 and box C/D, bulged and non-bulged strands, are listed in Table 2. Upon hybridisation, the Kt-7 and box C/D constructs had NC helices of 13 base pairs in length and C helices of 12 and 10 base pairs in length, respectively.

The principal of the FRET technique involves the transfer of energy from the donor to the acceptor fluorophore in relation to the distance between them, as discussed before. The emission spectrum of the donor overlaps with the excitation spectrum of the acceptor, hence the wavelengths of light emitted by the donor are absorbed by the acceptor which is excited, a process which is inversely proportional to the distance between the fluorophores [64], i.e. smaller the distance between the fluorophores, the greater the efficiency of fluorescence resonance energy transfer between them. As the E_{FRET} increases upon distance shortening the emission from the donor decreases due to energy being transferred to the acceptor whose emission increases as a result. Overall, in the context of k-turns, as the k-turn folds into the kinked conformation, the end-to-end distance between the 5'-terminally attached fluorophores decreases resulting in greater efficiency of fluorescence resonance energy transfer from the fluorescein donor to the Cy3 acceptor [62], which can be observed.

To calculate the $E_{FRET,}$ absorption spectra were first measured by re-suspending the ethanol precipitated hybrid RNA in 120µL 90mM Tris-borate (pH 8.4) and recording the absorbance in a 50mm path length cuvette using a NanoDrop 2000c spectrophotometer (Thermo Fisher Scientific). Fluorescein absorbs at the 490nm wavelength while Cy3 optimally absorbs at the 547nm wavelength although it has a broad absorption spectrum and also absorbs the 490nm wavelength of light. The overlapping fluorescein and Cy3 spectra were de-convoluted using a corresponding RNA species labelled only with Cy3 in order to obtain the fluorescein absorbance. Subsequently, the fluorophore absorption ratios, D_{490}/A_{547} and A_{490}/A_{547}, were calculated using a MATLAB program.

Secondly, fluorescence spectra were recorded for the hybrid RNA samples, with a concentration of 20nM in 500µL of 90mM Tris-borate (pH 8.4), at 4°C

using an SLM-Aminco 8100 fluorimeter. Spectra were corrected for lamp fluctuations and instrumental variations, and polarization artifacts were avoided by setting excitation and emission polarizers crossed at 54.7° as described in [108]. Spectra were obtained with both 490nm and 547nm excitation wavelengths for the bi-labelled hybrid RNA sample. At the 547nm wavelength only Cy3 is excited which subsequently emits at the 570nm wavelength. However, at the 490nm wavelength, both fluorescein and Cy3 are excited hence emit at the 520nm and 570nm wavelengths, respectively. Moreover, as well as the direct excitation of Cy3 by the 490nm wavelength, energy is also transferred to Cy3 from the excited fluorescein which also contributes towards the Cy3 emission. In order to extract the acceptor-only emission spectrum, the emission spectrum of a corresponding RNA species labelled only with the fluorescein was also obtained to be 'subtracted' from the emission spectrum of the bi-labelled RNA sample. As such, values of FRET efficiency were then measured using the acceptor normalization method [109] implemented in MATLAB, by inputting the corresponding D_{490}/A_{547} and A_{490}/A_{547} values obtained from the absorption spectra.

The E_{FRET} was measured in the steady-state after the addition of each Mg^{2+} ion concentration for a series of Mg^{2+} ion titrations (Table 3) performed on the k-turn RNA sample to observe it's folding at different Mg^{2+} ion concentrations. Subsequently, E_{FRET} as a function of Mg^{2+} ion concentration was analysed on the basis of a model in which the fraction of folded molecules corresponds to a simple two-state model for ion-induced folding [62], i.e.

$$E_{FRET} = E_0 + \Delta E_{FRET} \,.\, K_A\,[Mg^{2+}]^n/(1 + K_A\,[Mg^{2+}]^n)$$

where E_0 is the FRET efficiency of the RNA in the absence of added metal ions, ΔE_{FRET} is the increase in FRET efficiency at saturating metal ion concentration, $[Mg^{2+}]$ is the prevailing Mg^{2+} ion concentration, K_A is the apparent association constant for metal ion binding and n is a Hill coefficient. Data were fitted to this equation by nonlinear regression. The metal ion concentration at which the transition is half complete is given by $[Mg^{2+}]_{1/2} = (1/K_A)^{1/n}$ [62].

In some cases, after the Mg^{2+} ion titrations, the *Af*L7Ae protein was added to the hybrid RNA construct to a final concentration of 1μM, to observe the protein-induced folding ability. The $E_{FRET,}$ after the addition of the L7Ae protein, was measured as described above.

Titration	[$MgCl_2$ stock solution]	Volume of stock solution added (μL)	Final sample volume (μL)	Final [$MgCl_2$]
1	0.1mM	0.501	500.501	**0.1μM**
2	1mM	0.451	500.952	**1μM**
3	1mM	2.014	502.966	**5μM**
4	1mM	2.540	505.506	**10μM**
5	10mM	1.523	507.029	**40μM**
6	10mM	1.532	508.561	**70μM**
7	10mM	1.541	510.102	**100μM**
8	100mM	1.536	511.638	**400μM**
9	100mM	1.546	513.184	**700μM**
10	100mM	1.540	514.724	**1mM**
11	1M	2.069	516.793	**5mM**
12	1M	2.610	519.403	**10mM**
13	1M	21.870	541.273	**50mM**
14	1M	30.071	571.344	**100mM**

Table 8: Mg^{2+} ion titrations performed on the kink-turn RNA samples for analysis of metal ion-induced folding ability by FRET. The titrations were performed on the initial hybrid RNA sample with a concentration of 20nM in 500μL of 90mM Tris-borate (pH 8.4).

4.5 Comparative Gel Electrophoretic (CGE) Experiment

Kink-turn RNA folding in the presence of Mg^{2+} ions was also analysed by the comparative gel electrophoresis (CGE) experiment. The principle of the CGE experiment involves the observation of the relative mobility of the k-turn RNA in the electrophoresed gel which is dependent on the conformation it adopts according to the equation [67]:

$$\mu = \frac{Q}{\xi}\left\langle h_x^2 / L^2 \right\rangle$$

where μ is the mobility of the nucleic acid chain, Q is the effective charge, ξ is the frictional coefficient, h_x is the end-to-end length in the direction of the electric field and L is the contour length. Thus, if a RNA helix becomes kinked, i.e. the end-to-end distance is reduced; its electrophoretic mobility will be retarded. Therefore, the more folded the k-turn RNA is into the kinked conformation, the more retarded is its mobility in the gel. Gel electrophoresis is performed in native conditions, in the presence of Mg^{2+} ions, to observe the metal ion-induced folding ability of the k-turn RNA.

The DNA-RNA-DNA species for *H. marismortui* Kt-7 and its modified variants were chemically synthesised. The sequences of the unmodified, bulged and non-bulged, Kt-7 strands are listed in Table 2. The duplex species were prepared by annealing equimolar quantities of the appropriate oligonucleotides in 50mM Tris-HCl (pH 8.1) by slow cooling from 95°C to 4°C in a water bath. Thereafter, 2.5% ficoll was added to the hybridized species which were then electrophoresed in a 13% acrylamide: bis-acrylamide (29:1) gel in 90mM Tris.borate (pH 8.4) (1X TB). 2mM $MgCl_2$ was also added to the native gel. 200pmol of the hybrid RNA sample was loaded into one lane of the gel. After loading samples in the gel, electrophoresis was performed at 120V at 4°C for at least 72h, with recirculation of the 2mM $MgCl_2$-containing 1X TB buffer at >1 litre h^{-1}. The gel was stained using SYBR Gold (Life Technologies), washed in MilliQ water, and visualized on a Typhoon FLA 9500 (GE Healthcare). Gel electrophoresis was also performed in the presence of 2mM EDTA, in place of 2mM $MgCl_2$, in the gel and buffer. The components of the gel and buffer are summarised in Table 4.

	Component	Volume (*for pouring one gel*)
Gel (13% acrylamide: bis-Acrylamide (29:1)) (with 2mM $MgCl_2$/EDTA)	40% acrylamide: bis-acrylamide (29:1)	23mL
	10X TB	7mL
	MilliQ water	40mL
	1M $MgCl_2$ / 0.5M EDTA	140µL / 280µL
	10% APS	583µL
	TEMED	29.2µL
Buffer (1X TB) (with 2mM $MgCl_2$/EDTA)	10X TB	60mL
	MilliQ water	540mL
	1M $MgCl_2$ / 0.5M EDTA	1.2mL / 2.4mL

Table 9: The components of the gel and buffer, with the volumes required, for the execution of the gel electrophoresis experiment. The 13% acrylamide: bis-acrylamide (29:1) gel and 1X TB buffer contains 2mM $MgCl_2$ or 2mM EDTA for the native or denaturing gel electrophoresis experiment, respectively.

5. Results

5.1 The Sequence-Dependent Variance in the Folding Properties in Mg^{2+} Ions for the *Hm*Kt-7 and *Af*boxC/D Kink-Turns

The *Hm*Kt-7 k-turn sequence folds well in magnesium ions [51]. The *Af*boxC/D k-turn sequence, on the other hand, is unable to fold in magnesium ions [66]. The 3b:3n position in the k-turn sequence has been identified as a critical determinant that confers the ability of k-turns to undergo folding in metal ions. The *Af*boxC/D k-turn has the U:U base pair in the 3b:3n position, which in the *Hm*Kt-7 context allows moderate folding in metal ions. Yet the *Af*boxC/D k-turn exhibits no folding at all in metal ions [66]. This indicates the presence of other sequence determinants which are important for conferring metal ion- induced folding of k-turns. Comparison of the *Hm*Kt-7 and *Af*boxC/D k-turn sequences (Fig. 2) reveals three other positions that differ between the two, other than 3b:3n, which could potentially influence their folding properties in Mg^{2+} ions. These are the -1b:-1n and 4b:4n positions, as well as the loop. *Hm*Kt-7 has a C:G base pair in both -1b:-1n and 4b:4n positions, while *Af*boxC/D has a G:C base pair in those two positions. Also, *Hm*Kt-7 has a GAA loop while *Af*boxC/D has a CGU loop. Thus, these positions, alongside the 3b:3n position, have been investigated in this study for their effect on *Hm*Kt-7 and *Af*boxC/D k-turn folding properties in Mg^{2+} ions.

5.2 Systematic Modification of *Hm*Kt-7 and *Af*boxC/D Kink-Turn Sequences for Analysis of Folding in Mg^{2+} Ions by FRET

The folding of *Hm*Kt-7 and *Af* boxC/D k-turn sequences, and their variants, in response to the addition of Mg^{2+} ions has been investigated. From here on, the *Hm*Kt-7 and *Af* boxC/D k-turns will be named as the Kt-7 and box C/D k-turns, respectively. The Kt-7 and box C/D k-turn sequences were chemically synthesized, introducing modifications to the wild-type sequences as

required. Box C/D elements were incorporated into the Kt-7 sequence and vice versa. Initially, the -1b:-1n, 3b:3n, 4b:4n and loop positions were individually modified. Subsequently, double modifications were introduced into the k-turn sequences followed by triple modifications, until eventually all four sequence elements were exchanged between the Kt-7 and box C/D sequences converting Kt-7 into box C/D and vice versa.

The k-turn RNA synthesized was 5'-labelled with fluorophores, with the fluorescein donor on the bulged strand and Cy3 acceptor on the non-bulged strand, to allow the observation of k-turn folding in response to the addition of Mg^{2+} ions by the change in FRET efficiency. If the k-turn RNA folds into the kinked conformation upon the addition of Mg^{2+} ions or protein (added in some cases at the end of the magnesium ion titrations to observe protein-induced folding ability), the end-to-end distance between the fluorophores decreases resulting in greater efficiency of fluorescence resonance energy transfer from the donor to the acceptor. Hence, the emission from the fluorescein donor decreases as energy is transferred to the Cy3 acceptor whose emission increases correspondingly. The emission spectra, which are recorded in the steady-state after each magnesium ion titration, are then used to calculate the E_{FRET} using the acceptor normalization method [110]. Subsequently, E_{FRET} is plotted as a function of magnesium ion concentration and fitted to a two-state folding model (see Materials and Methods), in which an increase in E_{FRET} evidences an increase in the population of the folded k-turn species.

Initially, the wild-type Kt-7 and box C/D k-turns were analysed for folding in Mg^{2+} ions using FRET to confirm their folding and non-folding properties, respectively. The E_{FRET} data, as a function of Mg^{2+} ion concentration, for the unmodified Kt-7 and box C/D k-turns is plotted in Figures 11 and 12, respectively. After the titrations with the Mg^{2+} ions, 1μM L7Ae protein was added to the box C/D construct to observe its protein-induced folding ability. The E_{FRET} value obtained after the addition of L7Ae is also plotted in Figure 12.

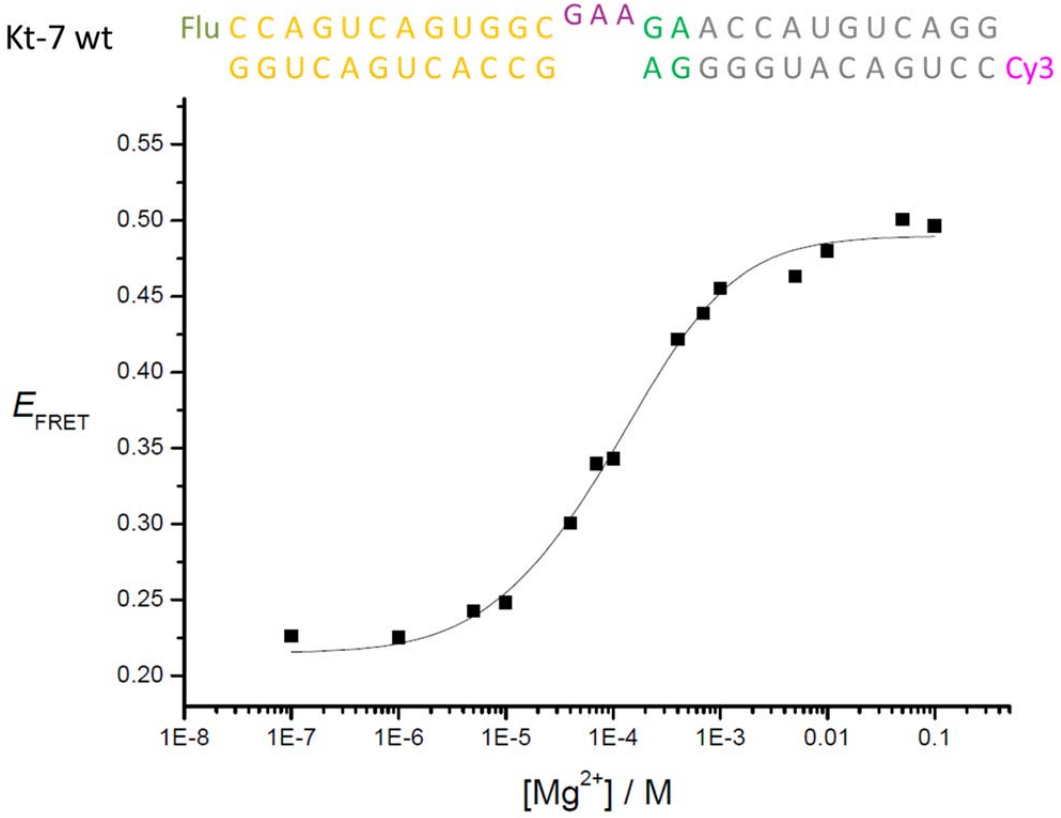


Figure 11: Mg^{2+} ion-dependent folding analysis of *H. marismortui* Kt-7 wild-type by FRET. The sequence of the wild-type Kt-7 construct used in the FRET experiment is displayed at the top. E_{FRET} (y-axis) has been plotted as a function of Mg^{2+} ion concentration (x-axis) and the curve fitted to the two-state model.

The unmodified Kt-7 folds well into the kinked conformation in the presence of Mg^{2+} ions shown by the increase in E_{FRET} upon titration with Mg^{2+} ions (Fig. 11), consistent with previous research [51]. The E_{FRET} data, as a function of Mg^{2+} ion concentration, is well fitted by the two-state folding model giving a $[Mg^{2+}]_{1/2}$ = 110 μM. The final E_{FRET} value is 0.49 corresponding to a well-folded k-turn. In this study, the final E_{FRET} value is taken as a measure of the extent of metal ion-dependant folding, i.e. the higher the final E_{FRET} value the better the folding of the k-turn. The complete set of initial and final E_{FRET} values for all k-turns studied in this work is tabulated in Table 1.

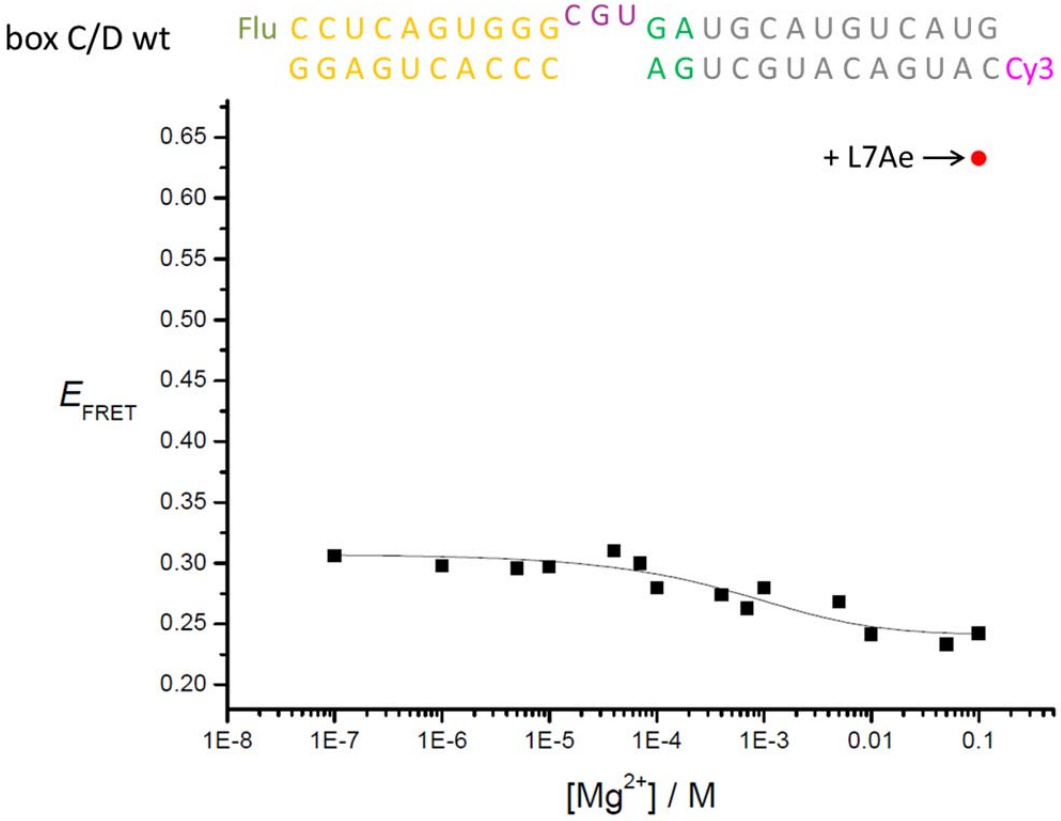


Figure 12: Mg^{2+} ion and L7Ae protein-dependent folding analysis of *A. fulgidus* box C/D wild-type by FRET. The sequence of the wild-type box C/D construct used in the FRET experiment is displayed at the top. E_{FRET} (y-axis) has been plotted as a function of Mg^{2+} ion concentration (x-axis) and the curve fitted to the two-state model. After the magnesium ion titrations, 1µM L7Ae was added and E_{FRET} calculated which is plotted as the red circle.

The unmodified box C/D k-turn is unable to fold into the kinked conformation in the presence of Mg^{2+} ions as depicted by the absence of increased E_{FRET} upon the addition of Mg^{2+} ions (Fig. 12), consistent with previous research [66]. Comparing the two data sets, the initial E_{FRET} value of box C/D of ~0.30 is higher than the corresponding value for Kt-7 of 0.21, suggesting a more kinked conformation for the unfolded state of the box C/D k-turn (discussed later). Nevertheless, upon the addition of Mg^{2+} ions, the E_{FRET} for box C/D subsequently decreases to a lower final E_{FRET} value of around 0.25. Generally, final E_{FRET} values below 0.30 represent unfolded k-turns while values above 0.30 are representative of foldable k-turns. Thus, box C/D is unable to fold in Mg^{2+} ions. However, after the Mg^{2+} ion titrations, the addition of 1µM L7Ae protein results in a big increase in the final E_{FRET} value for box C/D to 0.63 (Fig. 12, red circle), suggesting the complete folding of box C/D

into the k-turn conformation. This folding effect of L7Ae is obviously independent of the presence of Mg^{2+} ions [70] which have no folding effect. Thus, overall, the box C/D k-turn is able to fold into the kinked conformation when bound by the L7Ae protein, despite not being able to fold in the presence of Mg^{2+} ions alone.

Nevertheless, this variation in the folding properties in Mg^{2+} ions of Kt-7 and box C/D indicates the presence of sequence elements in the k-turns that determine the conformational response to the addition of metal ions. These sequence elements, namely the -1b:-1n, 3b:3n, 4b:4n and loop positions, have therefore been investigated for their effect on the folding properties of Kt-7 and box C/D in Mg^{2+} ions. The systematic modification of these positions in the Kt-7 and box C/D k-turn sequences and their effect on k-turn folding properties in Mg^{2+} ions is discussed in the following sections.

5.2.1 The -1b:-1n Position is the Key Sequence Determinant for Kink-Turn Folding Properties in Mg^{2+} Ions

The Kt-7 and box C/D k-turn sequences have C:G and G:C base pairs, respectively, in the -1b:-1n position. This sequence element was exchanged between the two k-turn sequences such that the modified Kt-7 and box C/D k-turns had G:C and C:G base pairs, respectively, in the -1b:-1n position. The folding properties in Mg^{2+} ions of the modified k-turns were then analysed by FRET. The E_{FRET} data, as a function of Mg^{2+} ion concentration, for the -1b:-1n modified Kt-7 and box C/D k-turns is plotted in Figures 13 and 14, respectively, alongside the E_{FRET} data for the corresponding unmodified sequences.

Subsequently, further modifications were introduced into the k-turn sequences at positions 3b:3n, 4b:4n and the loop, incorporating box C/D sequence elements into the Kt-7 k-turn sequence and vice versa. Initially, one further position was modified in addition to the -1b:-1n position (double modification), followed by the subsequent modification of a third position (triple modification) until eventually all four positions were modified,

converting Kt-7 to box C/D and vice versa. Folding in response to the addition of Mg^{2+} ions was analysed for each species using FRET. The final E_{FRET} values for all the Kt-7 and box C/D variant sequences analysed are plotted in Figure 15 in accordance to the identity of the -1b:-1n position. Note that while exchange of all four sequence elements effectively interconverts the two k-turn sequences, e.g. -1b:-1n=C:G + 3b:3n=A:G + 4b:4n=C:G + L=GAA modified box C/D has a similar sequence to the wild-type Kt-7, the final E_{FRET} values for these k-turns are not identical because of small differences in the lengths and sequences of the C and NC helices outside the core k-turn region. The complete sequence of the Kt-7 and box C/D k-turns used in the FRET experiments is presented in Materials and Methods.

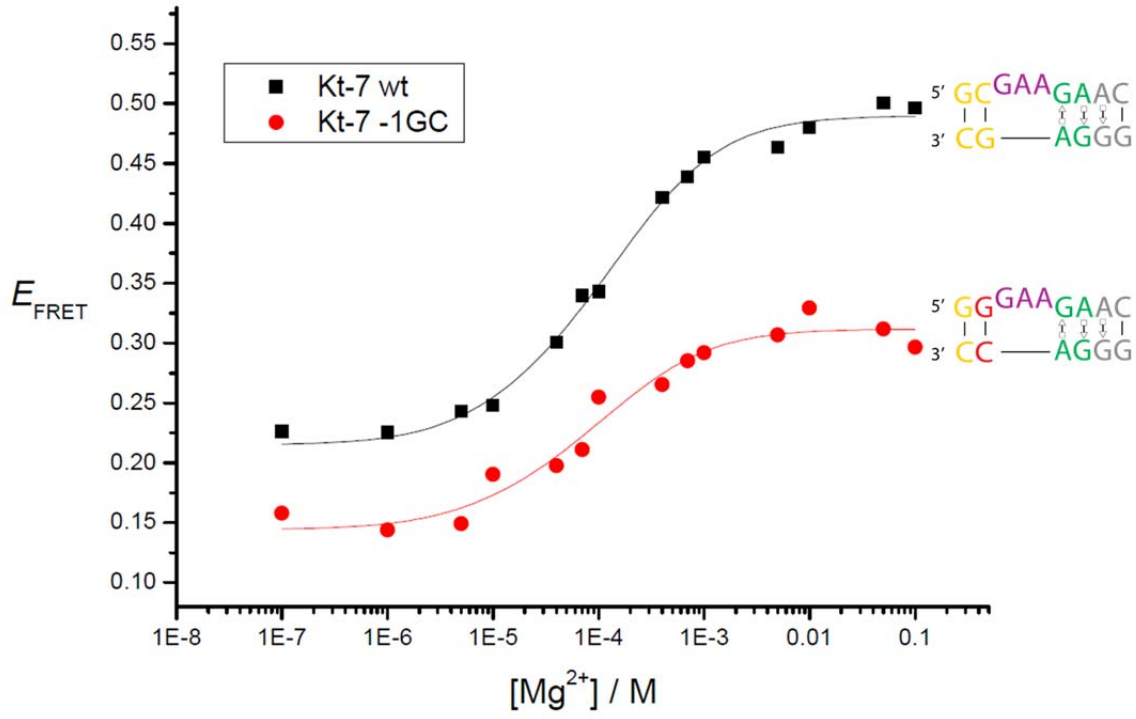


Figure 13: Mg^{2+} ion-dependent folding analysis of wild-type and -1b:1n=G:C modified *H. marismortui* Kt-7 by FRET. E_{FRET} (y-axis) has been plotted as a function of Mg^{2+} ion concentration (x-axis) and the curves for the wild-type Kt-7 (black squares) and -1b:-1n=G:C modified Kt-7 (red circles) are fitted to the two-state model.

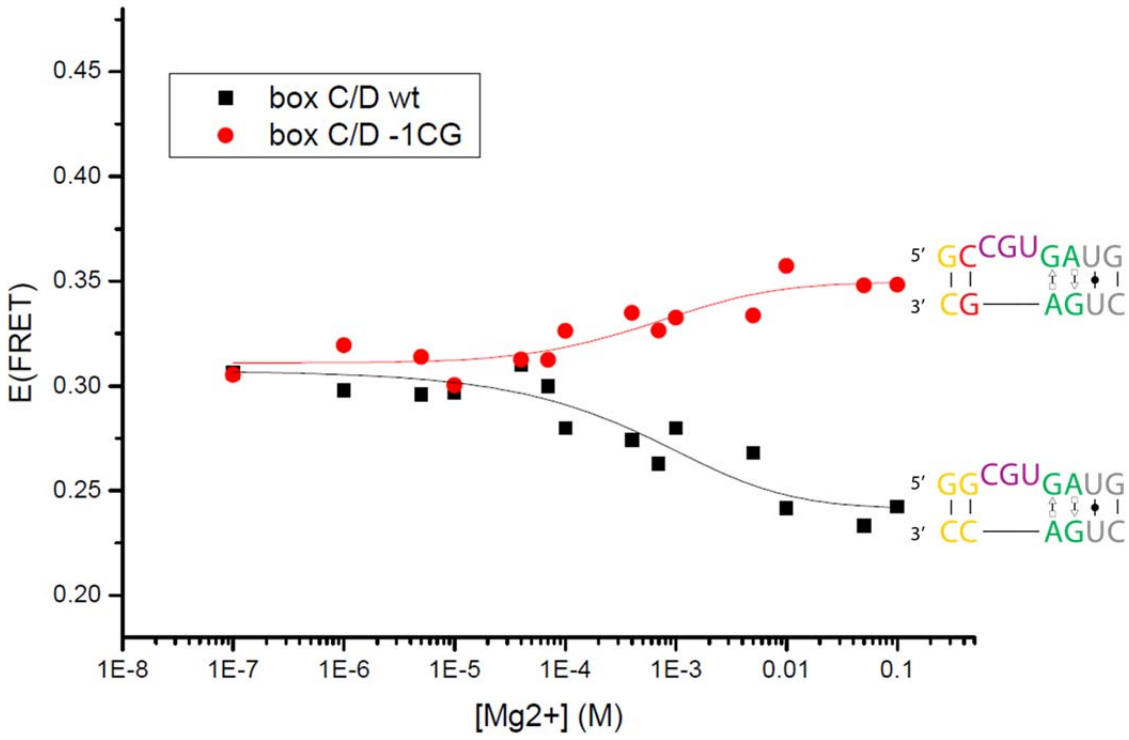


Figure 14: Mg^{2+} ion-dependent folding analysis of wild-type and -1b:-1n=C:G modified *A. fulgidus* box C/D by FRET. E_{FRET} (y-axis) has been plotted as a function of Mg^{2+} ion concentration (x-axis) and the curves for the wild-type box C/D (black squares) and -1b:-1n=C:G modified box C/D (red circles) are fitted to the two-state model.

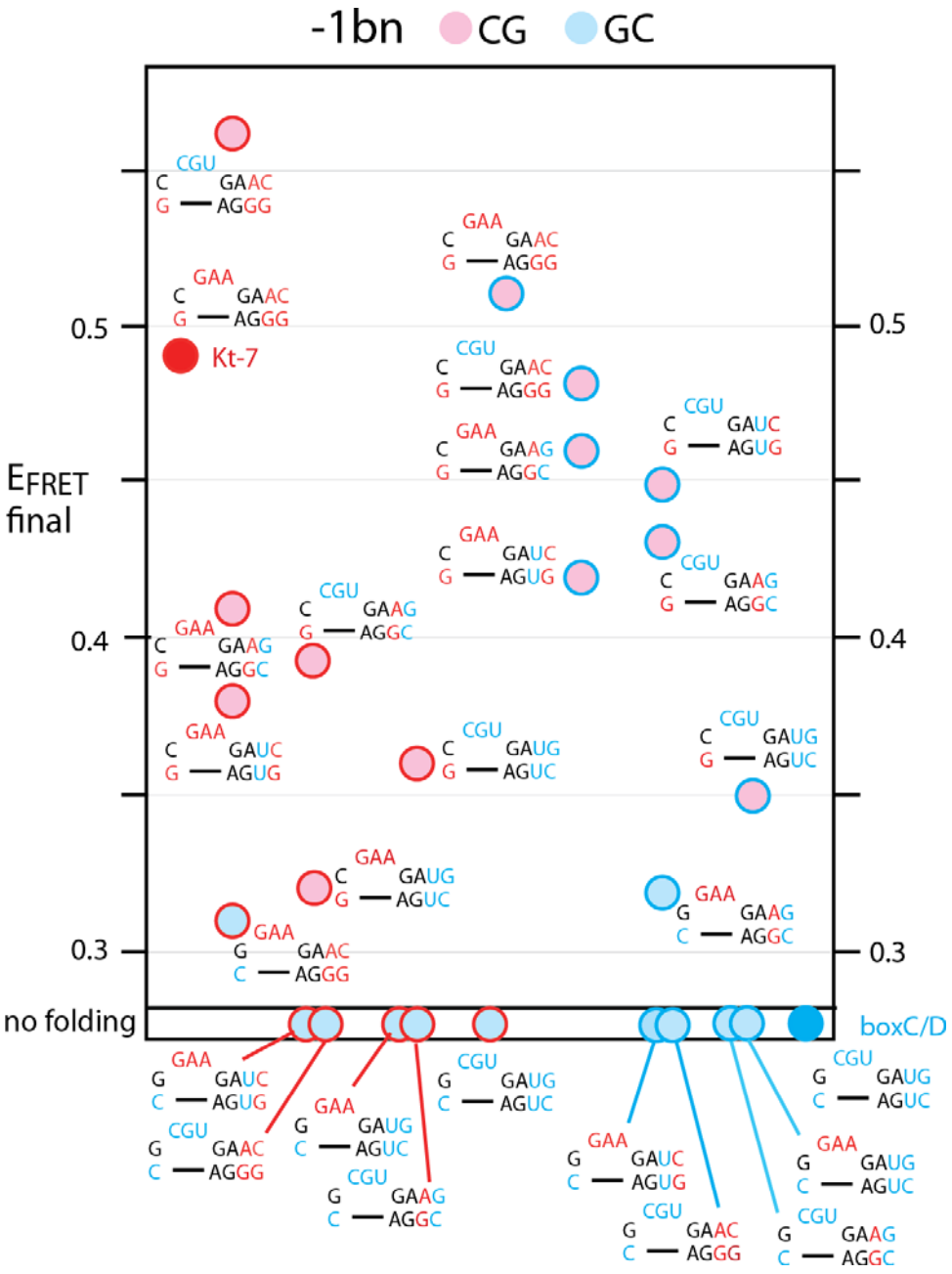


Figure 15: Plot of final E_{FRET} values in accordance to the identity of the -1b:-1n position of the kink-turn sequence for all Kt-7 and box C/D variant sequences analysed by FRET for folding properties in Mg^{2+} ions. The final E_{FRET} values for the Kt-7 and box C/D sequences are plotted on the left and right halves of the diagram, respectively. The unmodified Kt-7 and box C/D sequences are shown as red and blue filled circles to the extreme left and right sides of the diagram, respectively. The number of modifications in the Kt-7 sequence increase moving from the left to middle of the diagram. Similarly, the number of modifications in the box C/D sequence increase moving from right to middle of the diagram. Modified Kt-7 and box C/D sequences are represented by red and blue outlined circles, respectively, with -1b:-1n=C:G and -1b:-1n=G:C containing sequences represented by pink and light blue filled circles, respectively.

From Figure 13 it is apparent that modification of the -1b:-1n position in Kt-7 to G:C, the corresponding sequence element from box C/D, results in the impairment of k-turn folding in Mg^{2+} ions. The final E_{FRET} value is reduced from 0.49 to 0.31, converting Kt-7 from a good folding to a weakly folding k-turn in Mg^{2+} ions. This result clearly reveals the importance of the -1b:-1n position, whereby conversion from C:G to G:C can strongly impair the folding ability of the k-turn in Mg^{2+} ions. The likewise modification of the -1b:-1n position in box C/D to C:G results in an increase in E_{FRET} with Mg^{2+} ion concentration as opposed to the decrease in E_{FRET} observed in the wild-type box C/D species (Fig. 14). The resultant final E_{FRET} value is 0.35 showing that the modified box C/D k-turn undergoes a degree of folding in response to the addition of metal ions. Thus, modification of the -1b:-1n position from G:C to C:G can convert a non-folding k-turn into one with some folding ability in Mg^{2+} ions. Overall, the importance of the -1b:-1n position in k-turn folding in metal ions is clearly evident with -1b:-1n=C:G promoting the folding of k-turns in Mg^{2+} ions, while -1b:-1n=G:C represses the folding ability.

Interestingly, -1b:-1n=C:G is the only single modification in the box C/D k-turn that increases the final E_{FRET} value to above 0.30, i.e. induces the folding ability of the k-turn (Fig. 15). All other single modifications in box C/D still render the k-turn unable to fold. This suggests the -1b:-1n position to be the most important for determining folding characteristics of k-turns in Mg^{2+} ions compared to all other positions studied. Nevertheless, the modification of other positions in box C/D with Kt-7 sequence elements, in addition to the -1b:-1n position, further increases the final E_{FRET} value thus conferring better folding of the k-turn. This implies that the various positions have an additive effect on the k-turn folding properties in Mg^{2+} ions although the largest effect is exerted by the -1b:-1n position. For example, the box C/D construct with the -1b:-1n=C:G + 3b:3n=A:G + 4b:4n=C:G modifications has a higher final E_{FRET} value of 0.48, i.e. it fold wells into the k-turn conformation in Mg^{2+} ions, compared to the final E_{FRET} values of 0.43 and 0.35 of the weakly folded -1b:-1n=C:G + 3b:3n=A:G, and only -1b:-1n=C:G modified box C/D k-turns, respectively. This suggests that all of the different positions work additively to promote k-turn folding in Mg^{2+} ions. However, reverting the -1b:-1n=C:G

modification in this triply modified box C/D construct back to the wild-type -1b:-1n=G:C reduces the final E_{FRET} value to below 0.30, i.e. renders the k-turn unable to fold in Mg^{2+} ions, confirming the ultimate criticality of the -1b:-1n position as the major influencer of the folding properties of k-turns in Mg^{2+} ions. Furthermore, the box C/D construct with the highest final E_{FRET} value of 0.61 is modified at all four positions, -1b:-1n, 3b:3n, 4b:4n and the loop, with the corresponding Kt-7 sequence elements, thus is practically converted into Kt-7 and also behaves like Kt-7 by completely folding into the kinked conformation. This confirms the k-turn conformation inducing properties of the Kt-7 sequence elements, particularly -1b:-1n=C:G, 3b:3n=A:G and 4b:4n=C:G, revealing their additive effect. Generally, this shows the importance of the various positions in devising the k-turn folding properties in Mg^{2+} ions although the -1b:-1n position is the most critical.

In Kt-7, on the other hand, the single modification that results in the biggest reduction in the final E_{FRET} value is -1b:-1n=G:C (Fig. 15), again confirming this position to be the most important for influencing k-turn folding properties in Mg^{2+} ions. All single modifications in other positions result in a smaller reduction in the final E_{FRET} value. The further modification of other positions, alongside the -1b:-1n position, results in complete abolishment of k-turn folding ability in Mg^{2+} ions. This again demonstrates the additive effect of the various positions in influencing k-turn folding properties in metal ions. In general, all k-turn sequences that have a C:G in the -1b:-1n position have a final E_{FRET} value of ≥0.35, i.e. are folded into the kinked conformation in Mg^{2+} ions to some extent if not completely. On the other hand, all k-turn sequences that have a G:C in the -1b:-1n position have a final E_{FRET} value of ≤0.32. In this case, most of the k-turns are in the no folding zone with the exception of two k-turns, including the -1b:-1n=G:C modified Kt-7, which have very weak folding ability in Mg^{2+} ions. Thus, overall, the -1b:-1n=C:G sequence element of Kt-7 promotes the folding of k-turns in Mg^{2+} ions while the -1b:-1n=G:C sequence element of box C/D represses the folding of k-turns in Mg^{2+} ions. Ultimately, the -1b:-1n position is the most critical sequence determinant for k-turn folding properties in metal ions and is the major contributor amongst all positions studied.

5.2.1.1 The Exocyclic Amine of the G-1n Nucleobase is an Important Determinant for Kink-Turn Folding Properties in Mg^{2+} Ions

The preference of C:G over the G:C base pair in the -1b:-1n position has been established. The -1n position is known to form key cross-strand hydrogen bonds with the position 2b that stabilise the k-turn conformation [56]. The role of the G-1n nucleotide was therefore examined for its effect on Mg^{2+} ion-induced folding of the k-turn. The exocyclic amine group of the G-1n nucleobase was removed from Kt-7 by inosine (I) substitution at the -1n position, and the resultant k-turn was analysed for its folding properties in Mg^{2+} ions using FRET. The E_{FRET} data, as a function of Mg^{2+} ion concentration, for the G-1nI modified and unmodified Kt-7 k-turns is plotted in Figure 16.

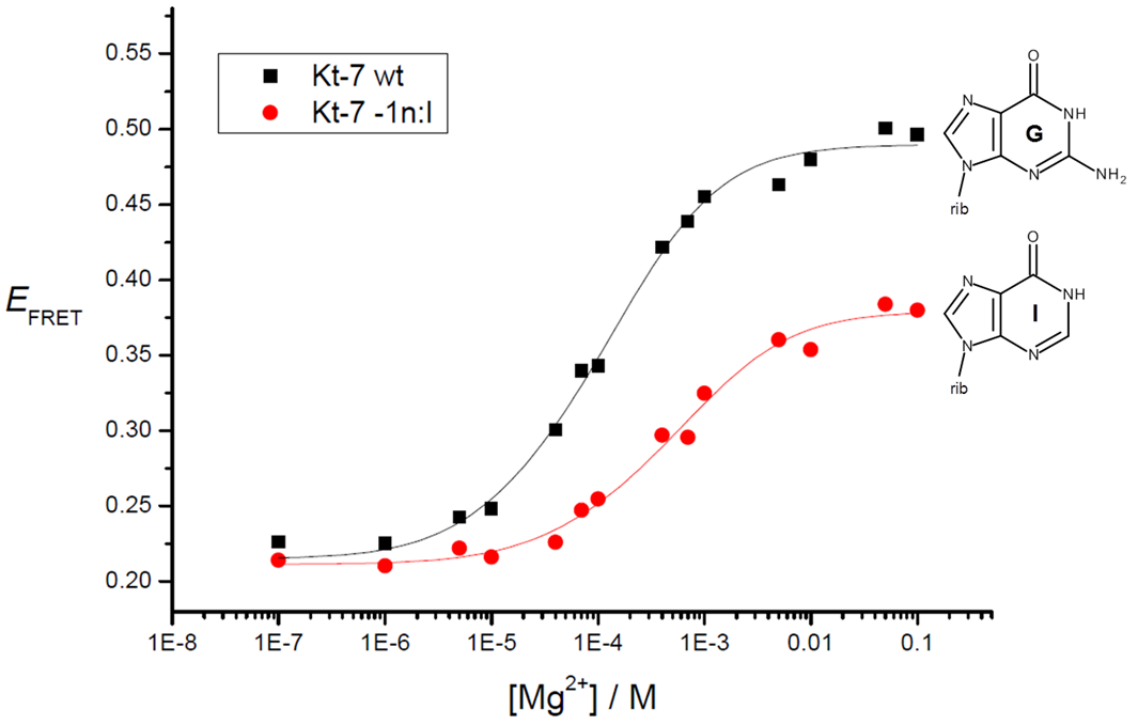


Figure 16: Mg^{2+} ion-dependent folding analysis of wild-type and G-1nI modified *H. marismortui* Kt-7 by FRET. E_{FRET} (y-axis) has been plotted as a function of Mg^{2+} ion concentration (x-axis) and the curves for the wild-type Kt-7 (black squares) and the G-1nI modified Kt-7 (red circles) are fitted to the two-state model. The structures of guanosine (G) and inosine (I) are displayed on the right.

The removal of the exocyclic amine group from the G-1n nucleobase by the substitution of guanosine in the -1n position of the Kt-7 k-turn sequence with inosine, results in the impairment of k-turn folding in Mg^{2+} ions. The final E_{FRET} value is reduced from 0.49 to 0.38, converting a well folding k-turn into one with intermediate folding ability in Mg^{2+} ions. This clearly reveals the importance of the exocyclic amine group in promoting the metal ion-induced folding of the k-turn possibly due to its involvement the k-turn stabilising A-minor interaction with the position 2b. Thus, the exocyclic amine group of the preferred G nucleobase in the -1n position of the k-turn sequence plays an important role in the stabilisation of the k-turn structure allowing it to completely fold into the kinked conformation in metal ions alone.

5.2.2 The Loop has a Negligible Effect on Kink-Turn Folding Properties in Mg^{2+} Ions

The loop sequences are GAA and CGU in the Kt-7 and box C/D k-turns, respectively. The Kt-7 and box C/D loops were exchanged such that the modified Kt-7 and box C/D k-turns had the CGU and GAA loops, respectively. These modified k-turns were then analysed for their folding properties in Mg^{2+} ions by FRET. The E_{FRET} data, as a function of Mg^{2+} ion concentration, for the loop modified Kt-7 and box C/D k-turns is plotted in Figures 16 and 17, respectively, alongside the E_{FRET} data for the corresponding unmodified sequences. The subsequent modification of other positions in the Kt-7 and box C/D k-turn sequences for analysis of their effect on k-turn folding properties in Mg^{2+} ions by FRET, has revealed final E_{FRET} values which are plotted in Figure 18 according to the identity of the loop sequence.

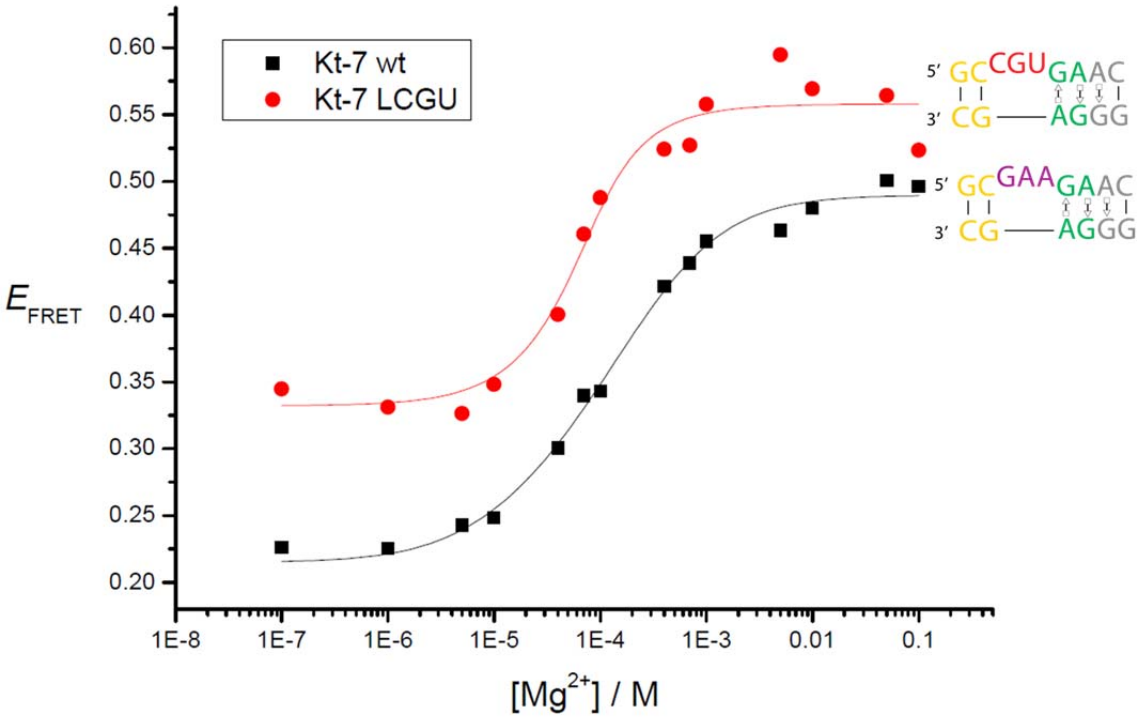


Figure 17: Mg^{2+} ion-dependent folding analysis of wild-type and L=CGU modified *H. marismortui* Kt-7 by FRET. E_{FRET} (y-axis) has been plotted as a function of Mg^{2+} ion concentration (x-axis) and the curves for the wild-type Kt-7 (black squares) and L=CGU modified Kt-7 (red circles) are fitted to the two-state model.

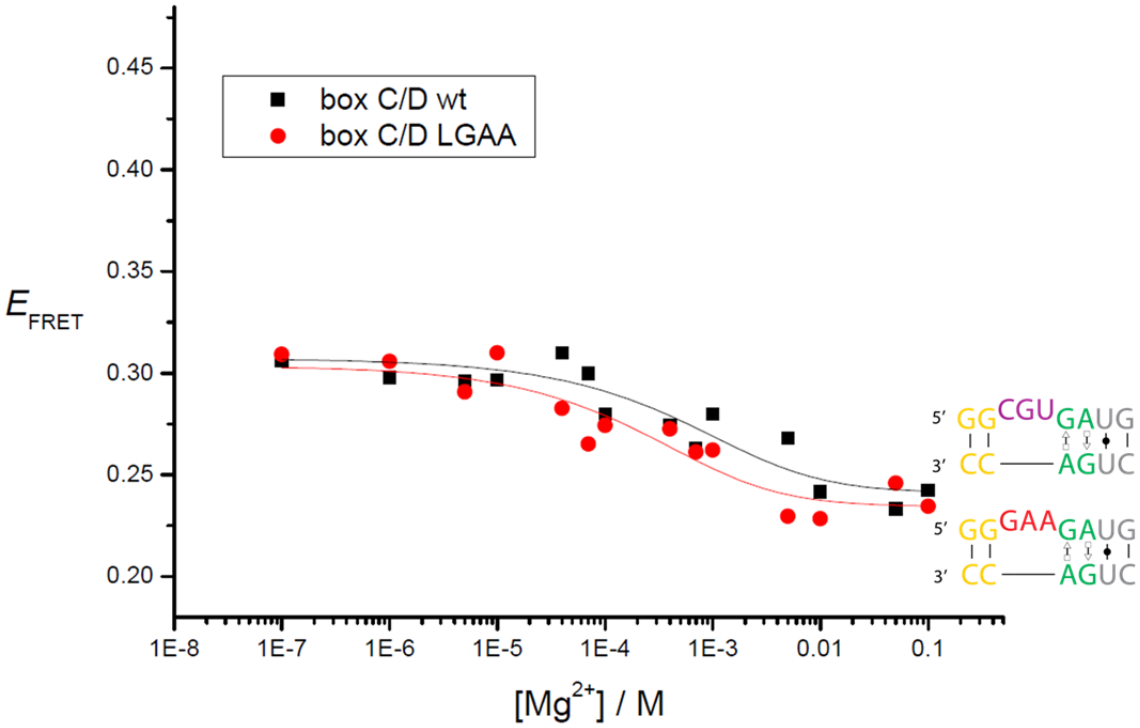


Figure 18: Mg^{2+} ion-dependent folding analysis of wild-type and L=GAA modified *A. fulgidus* box C/D by FRET. E_{FRET} (y-axis) has been plotted as a function of Mg^{2+} ion concentration (x-axis) and the curves for the wild-type box C/D (black squares) and L=GAA modified box C/D (red circles) are fitted to the two-state model.

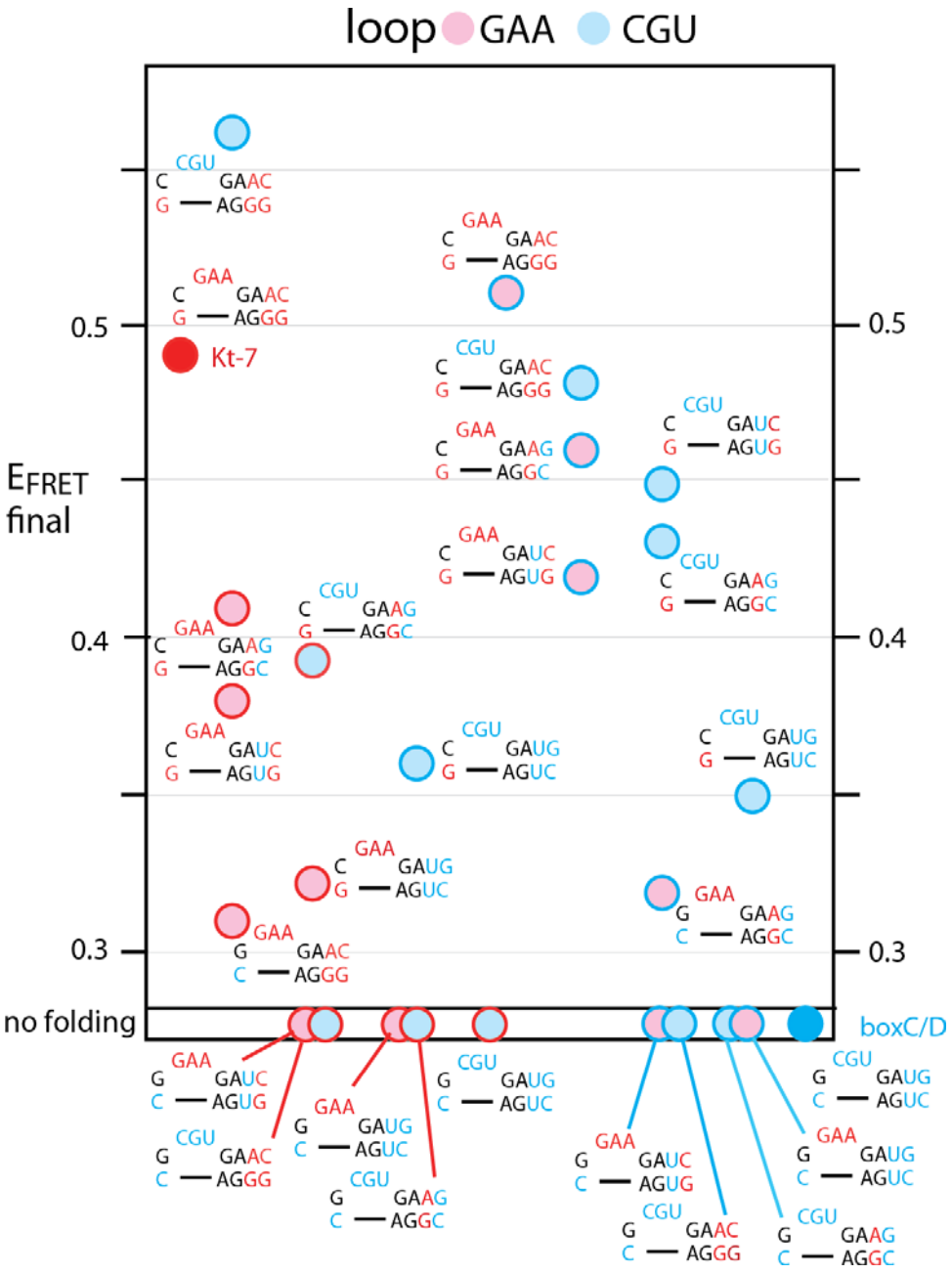


Figure 19: Plot of final E_{FRET} values in accordance to the identity of the loop sequence for all Kt-7 and box C/D variant sequences analysed by FRET for folding properties in Mg^{2+} ions. The final E_{FRET} values for the Kt-7 and box C/D sequences are plotted on the left and right halves of the diagram, respectively. The unmodified Kt-7 and box C/D sequences are shown as red and blue filled circles to the extreme left and right sides of the diagram, respectively. The number of modifications in the Kt-7 sequence increase moving from the left to middle of the diagram. Similarly, the number of modifications in the box C/D sequence increase moving from right to middle of the diagram. Modified Kt-7 and box C/D sequences are represented by red and blue outlined circles, respectively, with L=GAA and L=CGU containing sequences represented by pink and light blue filled circles, respectively.

Modification of Kt-7 with the CGU loop of box C/D results in the higher final E_{FRET} value of 0.56 compared to 0.49 of the unmodified Kt-7 (Fig. 16). Interestingly, this is the only single modification in Kt-7 that results in the increase of the final E_{FRET} value implying the enhancement of the folding ability of Kt-7 in Mg^{2+} ions. Nevertheless, the CGU loop also results in the relatively higher E_{FRET} value of 0.33 for the initial state of Kt-7 prior to the addition of Mg^{2+} ions, compared to the 0.21 initial E_{FRET} value for the unmodified Kt-7. This is consistent with the initial E_{FRET} value of around 0.30 for the unmodified box C/D k-turn which also has a CGU loop. Overall, this suggests that the unfolded k-turn structure, which is assumed to behave like a simple three nucleotide bulge in a duplex RNA molecule, is more bent with the CGU loop sequence than with the GAA loop sequence. It has been previously reported that the extent of kinking by base bulges in double-stranded RNA is dependent on the sequence of the bulge [111]. The new data indicates that the CGU loop results in greater kinking of the RNA molecule in the unfolded state in the absence of Mg^{2+} ions, as well as in the folded state in the presence of Mg^{2+} ions as observed in Kt-7.

Further exchange of positions in Kt-7, in addition to the loop, with sequence elements from box C/D, result in lower final E_{FRET} values (Fig. 18) suggesting the impairment of metal ion-dependant k-turn folding. In opposition to the enhancing properties of the CGU loop, this confirms the repressive contribution of the box C/D sequence elements, -1b:-1n=G:C, 3b:3n=U:U and 4b:4n=G:C, towards k-turn folding properties in Mg^{2+} ions, which work additively. For example, the subsequent modification of the 4b:4n position with the G:C base pair in the L=CGU modified Kt-7, results in the reduction of the final E_{FRET} value from 0.56 to 0.39. This final E_{FRET} value is further reduced to 0.36 when the 3b:3n position is additionally modified to U:U suggesting the additive effect of the various positions on k-turn folding properties in Mg^{2+} ions. Nevertheless, the -1b:-1n position exerts the greatest effect since only the -1b:-1n=G:C modification can completely abolish the folding of the L=CGU modified Kt-7 as dictated by the final E_{FRET} value of <0.30. Thus, modification of the -1b:-1n position from the C:G to the G:C base pair can convert a completely well folded k-turn into a non-foldable one

suggesting the utmost criticality of this sequence element in determining the folding properties of k-turns in metal ions.

The enhancing effect of the CGU loop on k-turn folding in Mg^{2+} ions, however, is not apparent in box C/D which is unable to fold with both the CGU and GAA loops (Fig. 17). Interestingly, the 3b:3n=A:G modification promotes the folding of the GAA loop-containing box C/D k-turn although to a very weak extent as depicted by the final E_{FRET} value of 0.32. However, the same modification still renders the CGU loop-containing box C/D k-turn unable to fold in Mg^{2+} ions with a final E_{FRET} value of <0.30. This suggests the CGU loop to have a repressive effect on k-turn folding in Mg^{2+} ions, while the GAA loop has a promotive effect. Similarly, the most well folded box C/D k-turn is modified with all four Kt-7 sequence elements, -1b:-1n=C:G, 3b:3n=A:G, 4b:4n=C:G as well as L=GAA. In this construct, the reversion of the loop back to the wild-type CGU loop results in the reduction of the final E_{FRET} value from 0.51 to 0.48, i.e. kink-turn folding is impaired to a slight extent. This again implies the preference of the GAA loop over the CGU loop to promote folding of the k-turn.

Overall, the loop can have both promotive and repressive effects on k-turn folding in metal ions although the magnitude of effect is rather small and overwhelmed by the effect exerted by the other positions studied. Both CGU and GAA loop-containing k-turn sequences are found spanning the entire range of final E_{FRET} values (Fig. 18), i.e. both loop sequences are present in well folded k-turns as well as in non-foldable ones, suggesting that the loop sequence is not important for influencing the k-turn folding properties in Mg^{2+} ions and that its effect is rather negligible.

5.2.3 The 3b:3n and 4b:4n Positions have Intermediate Importance in Conferring Kink-Turn Folding Properties in Mg^{2+} Ions

Kt-7 has the 3b:3n=A:G and 4b:4n=C:G sequence elements, while box C/D has the 3b:3n=U:U and 4b:4n=G:C sequence elements. These sequence

elements were individually exchanged between the two k-turn sequences such that the modified Kt-7 k-turn had either the 3b:3n=U:U or 4b:4n=G:C sequence element, while the modified box C/D k-turn had either the 3b:3n=A:G or 4b:4n=C:G sequence element. These modified Kt-7 and box C/D k-turns were then analysed for their folding properties in Mg^{2+} ions by FRET. The E_{FRET} data, as a function of Mg^{2+} ion concentration, for the 3b:3n=U:U / 4b:4n=G:C modified Kt-7 and 3b:3n=A:G / 4b:4n=C:G modified box C/D k-turns is plotted in Figures 19 and 20, respectively, alongside the E_{FRET} data for the corresponding unmodified sequences. Subsequently, other positions were also modified in these k-turns to determine their effect on the k-turn folding properties in Mg^{2+} ions. All the variant Kt-7 and box C/D k-turn sequences were analyzed by FRET and the final E_{FRET} values obtained are plotted in Figures 21 and 22 according to the identity of the 3b:3n and 4b:4n positions, respectively.

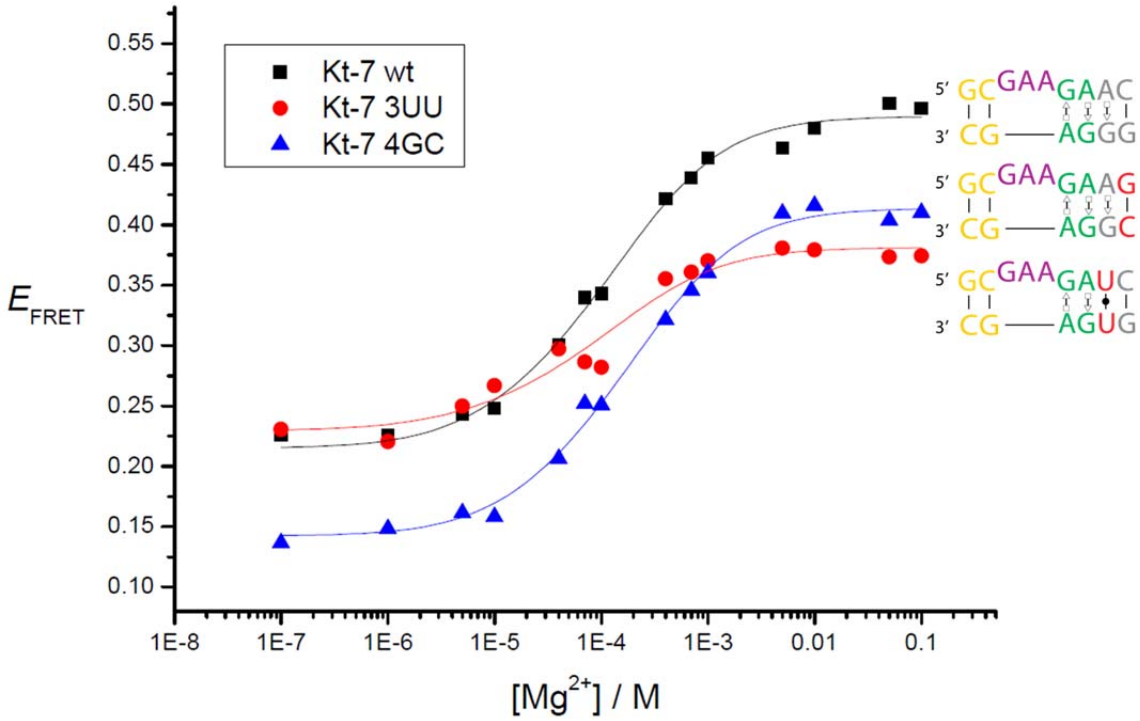


Figure 20: Mg^{2+} ion-dependent folding analysis of wild-type and 3b:3n=U:U / 4b:4n=G:C modified *H. marismortui* Kt-7 by FRET. E_{FRET} (y-axis) has been plotted as a function of Mg^{2+} ion concentration (x-axis) and the curves for the wild-type Kt-7 (black squares), 3b:3n=U:U modified Kt-7 (red circles) and 4b:4n=G:C modified Kt-7 (blue triangles) are fitted to the two-state model.

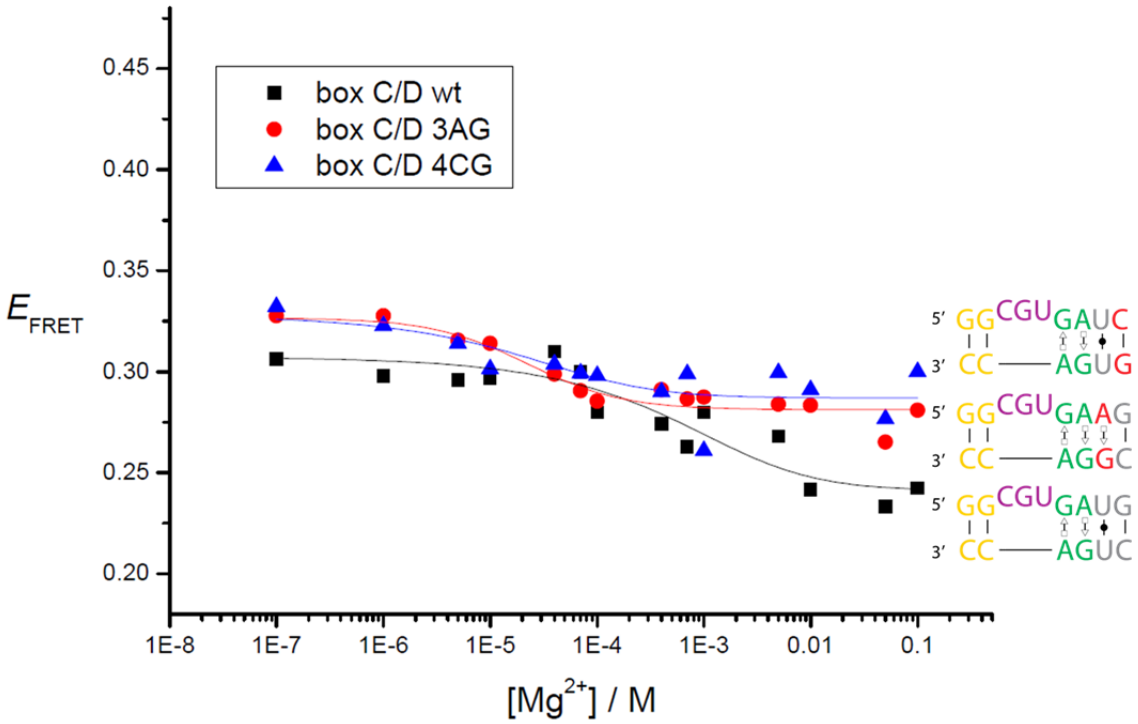


Figure 21: Mg^{2+} ion-dependent folding analysis of wild-type and 3b:3n=A:G / 4b:4n=C:G modified *A. fulgidus* box C/D by FRET. E_{FRET} (y-axis) has been plotted as a function of Mg^{2+} ion concentration (x-axis) and the curves for the wild-type box C/D (black squares), 3b:3n=A:G modified box C/D (red circles) and 4b:4n=C:G modified box C/D (blue triangles) are fitted to the two-state model.

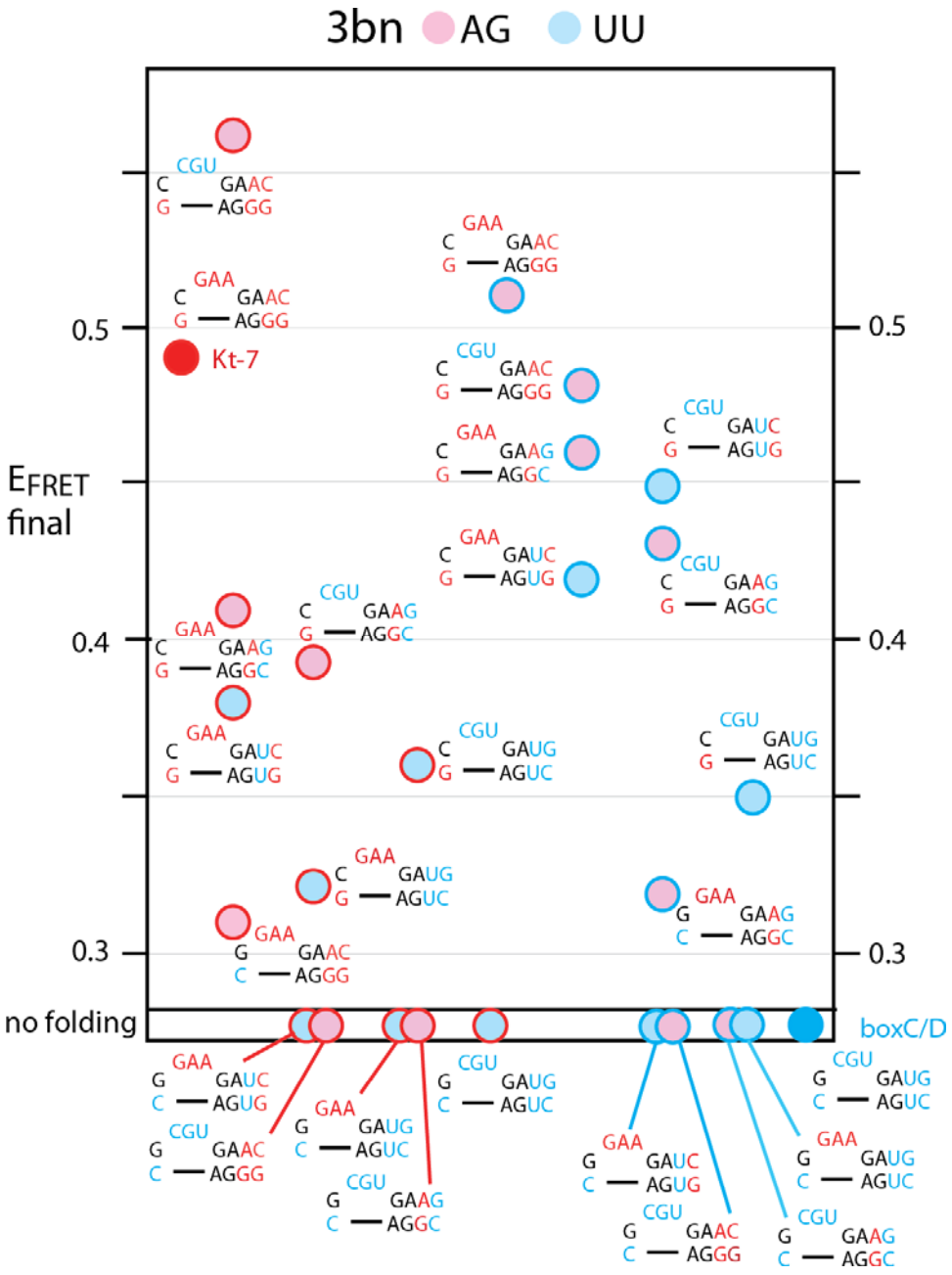


Figure 22: Plot of final E_{FRET} values in accordance to the identity of the 3b:3n position of the kink-turn sequence for all Kt-7 and box C/D variant sequences analysed by FRET for folding properties in Mg^{2+} ions. The final E_{FRET} values for the Kt-7 and box C/D sequences are plotted on the left and right halves of the diagram, respectively. The unmodified Kt-7 and box C/D sequences are shown as red and blue filled circles to the extreme left and right sides of the diagram, respectively. The number of modifications in the Kt-7 sequence increase moving from the left to middle of the diagram. Similarly, the number of modifications in the box C/D sequence increase moving from right to middle of the diagram. Modified Kt-7 and box C/D sequences are represented by red and blue outlined circles, respectively, with 3b:3n=A:G and 3b:3n=U:U containing sequences represented by pink and light blue filled circles, respectively.

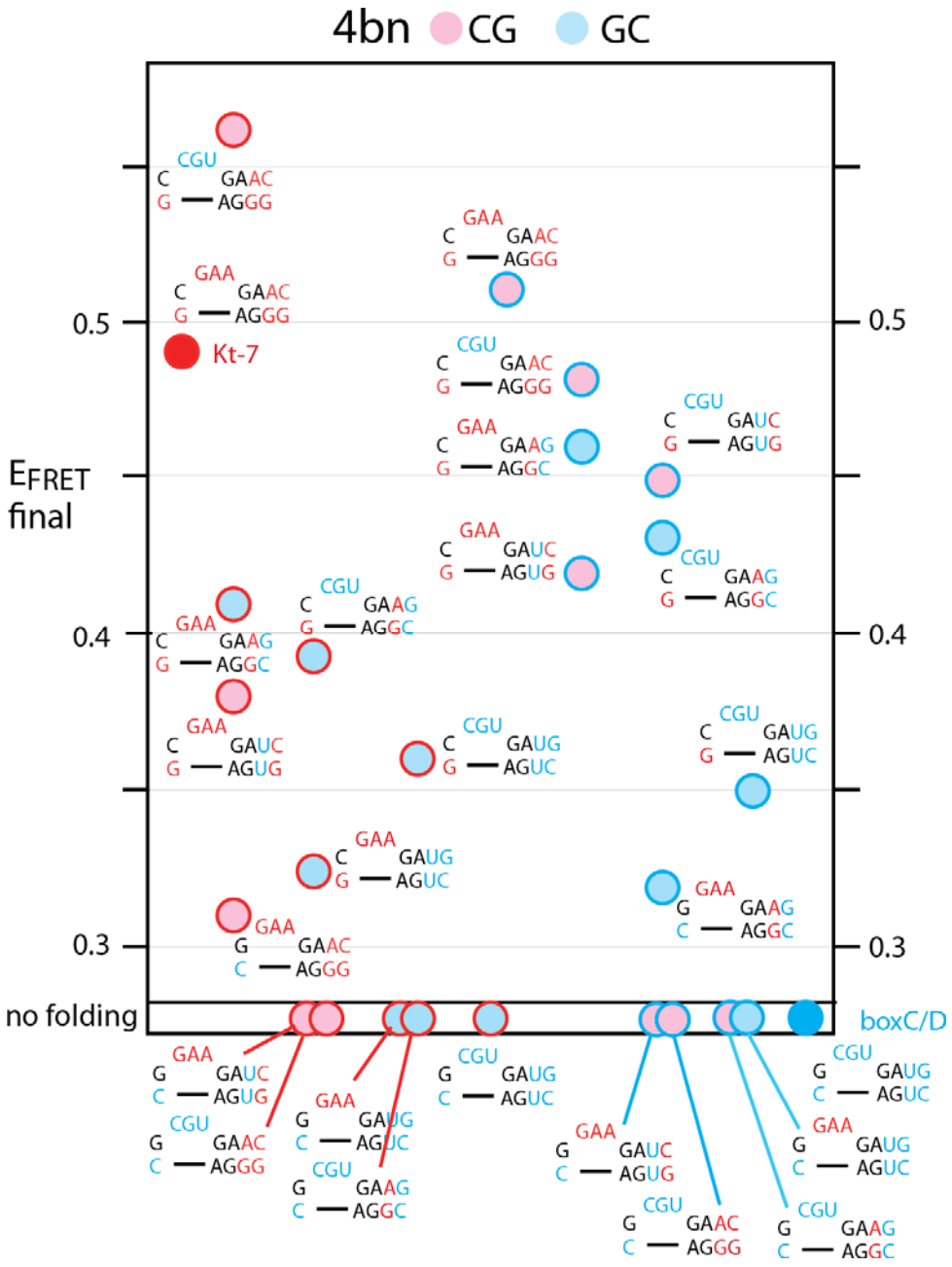


Figure 23: Plot of final E_{FRET} values in accordance to the identity of the 4b:4n position of the kink-turn sequence for all Kt-7 and box C/D variant sequences analysed by FRET for folding properties in Mg^{2+} ions. The final E_{FRET} values for the Kt-7 and box C/D sequences are plotted on the left and right halves of the diagram, respectively. The unmodified Kt-7 and box C/D sequences are shown as red and blue filled circles to the extreme left and right sides of the diagram, respectively. The number of modifications in the Kt-7 sequence increase moving from the left to middle of the diagram. Similarly, the number of modifications in the box C/D sequence increase moving from right to middle of the diagram. Modified Kt-7 and box C/D sequences are represented by red and blue outlined circles, respectively, with 4b:4n=C:G and 4b:4n=C:G containing sequences represented by pink and light blue filled circles, respectively.

Modification of the Kt-7 sequence with the 3b:3n=U:U sequence element of box C/D reduces the final E_{FRET} value from 0.49 to 0.38, while modification with the 4b:4n=G:C sequence element of box C/D reduces the final E_{FRET} value to 0.41 (Fig.19). Thus, inclusion of the 3b:3n=U:U and 4b:4n=G:C sequence elements convert a well folding k-turn into one with intermediate folding ability in Mg^{2+} ions suggesting their repressive effect on k-turn folding properties in metal ions. In combination, the 3b:3n=U:U and 4b:4n=G:C sequence modifications further reduce the final E_{FRET} value of Kt-7 to 0.32 (Fig. 21 and 22) resulting in even weaker folding of the k-turn in Mg^{2+} ions. This suggests that the 3b:3n and 4b:4n positions work additively to influence k-turn folding properties in Mg^{2+} ions.

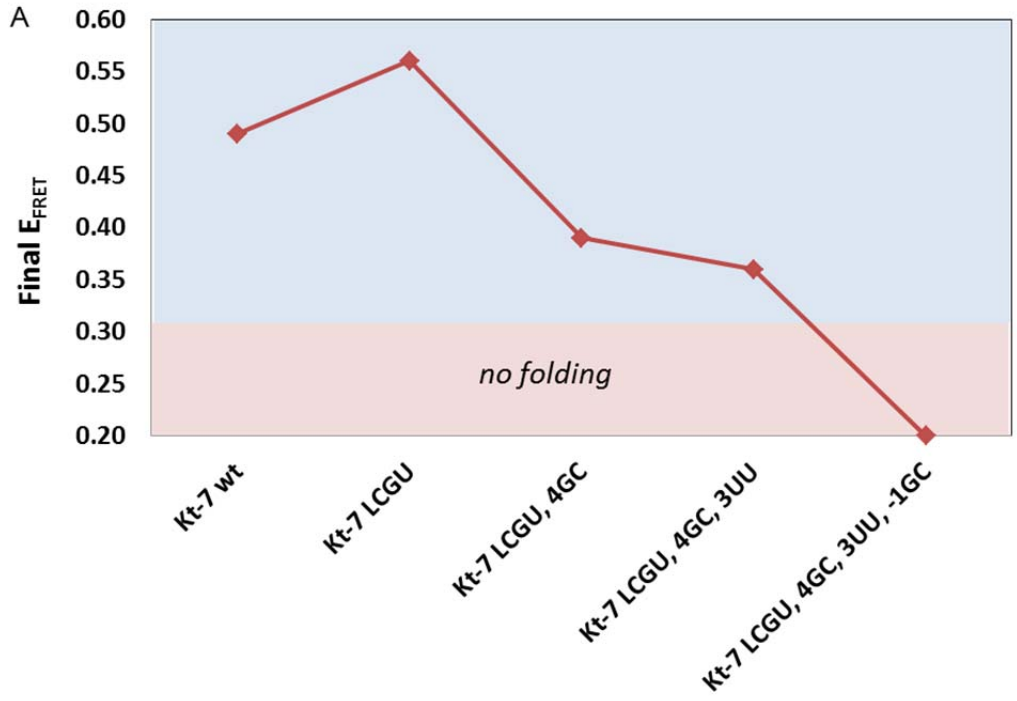
A further example of the additive effect on the k-turn folding properties in Mg^{2+} ions is revealed by the subsequent modification of the 3b:3n=U:U + 4b:4n=G:C modified Kt-7 sequence with the -1b:-1n=G:C sequence element which results in complete abolishment of the k-turn folding ability in Mg^{2+} ions as depicted by the final E_{FRET} value of <0.30 (Fig. 21 and 22). The singly modified Kt-7 with the -1b:-1n=G:C sequence element, on the other hand, is still able to weakly fold in Mg^{2+} ions with a final E_{FRET} value of 0.31, confirming that the -1b:-1n position works in addition to positions 3b:3n and 4b:4n to prevent folding of the triply modified Kt-7 k-turn. Thus, the various positions have an additive effect on k-turn folding properties in Mg^{2+} ions, although the strongest effect is exerted by the -1b:-1n position. This is further shown by the incorporation of the -1b:-1n=G:C sequence element into the 3b:3n=U:U modified Kt-7 and the 4b:4n=G:C modified Kt-7 sequences which also result in k-turns with no folding ability in Mg^{2+} ions as revealed by the final E_{FRET} values of <0.30. This implies that the folding enhancement effect of the 4b:4n=C:G and 3b:3n=A:G sequence elements in the two respective k-turns is overwhelmed by the greater repressive effect of the -1b:-1n=G:C sequence element. Therefore, the -1b:-1n position can be regarded as being more important than the 3b:3n and 4b:4n positions in influencing k-turn folding properties in Mg^{2+} ions. Nevertheless, the 3b:3n and 4b:4n positions also affect the k-turn folding properties in metal ions, although to a lesser extent than the -1b:-1n position.

Similarly, the enhancing effect of the 3b:3n=A:G and 4b:4n=C:G sequence elements is not apparent in box C/D as the incorporation of these sequence elements into box C/D still render the k-turns unable to fold in Mg^{2+} ions as depicted by the decrease in E_{FRET} upon titration with Mg^{2+} ions to the final E_{FRET} values of <0.30 (Fig. 20). Again, this is due to the presence of the wild-type -1b:-1n=G:C sequence element in the box C/D k-turn sequence whose repressive effect outweighs the promotive effects of the 3b:3n=A:G and 4b:4n=C:G sequence elements. Nevertheless, in the -1b:-1n=G:C modified box C/D k-turn, the incorporation of the 3b:3n=A:G or 4b:4n=C:G sequence element increases the final E_{FRET} value from 0.35 to 0.43 or 0.45, respectively (Fig. 21 and 22). Thus, the weakly folding k-turn is converted into one with intermediate folding ability in Mg^{2+} ions, confirming the promotive effect of the 3b:3n=A:G and 4b:4n=C:G sequence elements on k-turn folding in Mg^{2+} ions. Combined together these modifications further increase the final E_{FRET} value of the triply modified box C/D k-turn to 0.48, suggesting the better folding of the k-turn. This again confirms the additive effect of the various positions on k-turn folding properties in Mg^{2+} ions.

On the whole, the 3b:3n=A:G and 4b:4n=C:G sequence element-containing k-turns span the entire range of final E_{FRET} values (Fig. 21 and 22), i.e. the 3b:3n=A:G and 4b:4n=C:G sequence elements are present in both well folded and non-foldable k-turns. The 3b:3n=U:U and 4b:4n=G:C sequence elements, however, are only present in k-turns with intermediate or non-folding ability in Mg^{2+} ions. This suggests that for the complete folding of the k-turn the 3b:3n=A:G and 4b:4n=C:G sequence elements are necessary, suggesting their importance in promoting Mg^{2+} ion-induced folding of the k-turn. However, the presence of the 3b:3n=A:G and 4b:4n=C:G sequence elements by themselves does not guarantee k-turn folding in Mg^{2+} ions suggesting their magnitude of effect to be comparatively lower than of the utmost critical -1b:-1n sequence element. Hence, the 3b:3n and 4b:4n positions can be regarded as having intermediate importance in determination of the k-turn folding properties in metal ions.

5.2.4 The Transition from Kt-7 to box C/D and vice versa: A Summary of the Additive Effect of the Various Sequence Elements

The Kt-7 k-turn was converted into the box C/D k-turn, and vice versa, by the incorporation of the four box C/D sequence elements, -1b:-1n=G:C, 3b:3n=U:U, 4b:4n=G:C and L=CGU, and the four Kt-7 sequence elements, --1b:-1n=C:G, 3b:3n=A:G, 4b:4n=C:G and L=GAA, into the Kt-7 and box C/D k-turn sequences, respectively. The sequence elements were introduced one at a time in the order of least to most significant inKt-7 and vice versa in box C/D. Each modified Kt-7 and box C/D k-turn was analysed for its folding properties in Mg^{2+} ions using FRET. The resulting final E_{FRET} values for each modified Kt-7 and box C/D k-turn are plotted in Figures 23A and 24A, respectively, along with some representative plots of E_{FRET} as a function of Mg^{2+} ion concentration (Figures 23B and 24B). Furthermore, the final E_{FRET} values for all the variant Kt-7 and box C/D k-turn sequences analyzed by FRET are plotted in Figure 25.



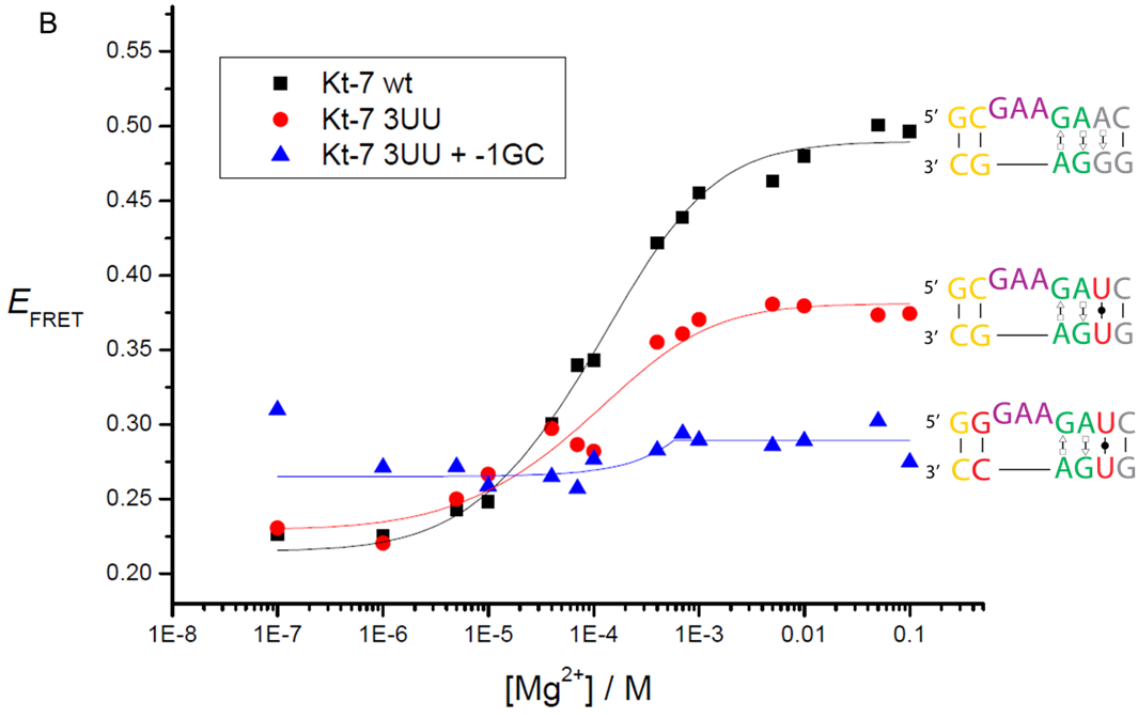
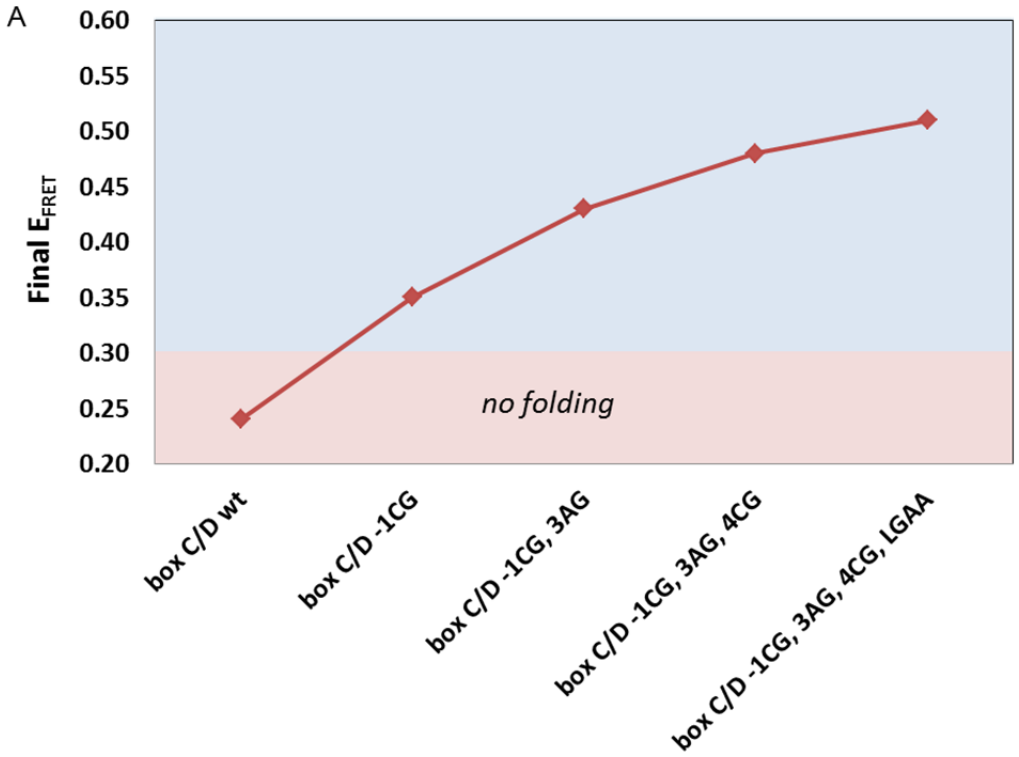


Figure 24: Transition from the Kt-7 to the box C/D kink-turn sequence. A. A plot of the final E_{FRET} values for the Kt-7 k-turn sequences subsequently modified with the box C/D sequence elements in the order of least to most significant, i.e. L=CGU, 4b:4n=G:C, 3b:3n=U:U and -1b:-1n=G:C. B. E_{FRET} (y-axis) has been plotted as a function of Mg^{2+} ion concentration (x-axis) and the curves fitted to the two-state model for the wild-type Kt-7 (black squares), 3b:3n=U:U modified Kt-7 (red circles) and 3b:3n=U:U + -1b:-1n=G:C modified Kt-7 (blue triangles).


A
0.60
0.55
0.50
0.45
0.40
0.35
0.30
0.25
0.20
Final E_{FRET}
no folding
box C/D wt
box C/D -1CG
box C/D -1CG, 3AG
box C/D -1CG, 3AG, 4CG
box C/D -1CG, 3AG, 4CG, LGAA

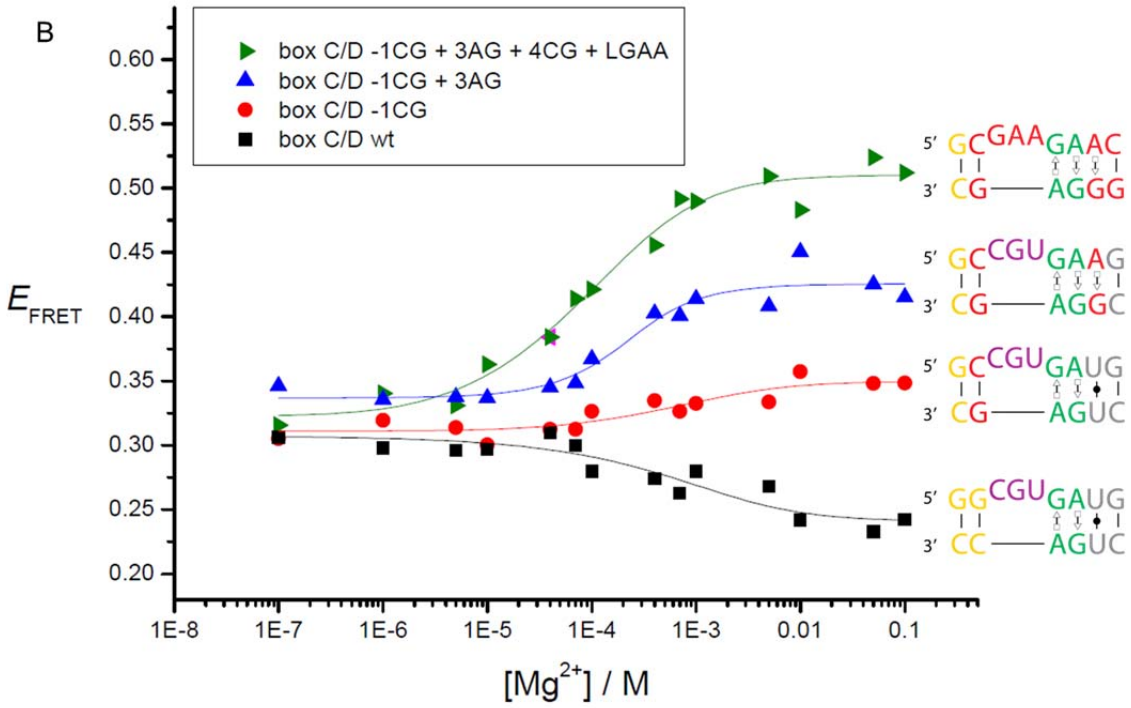


Figure 25: Transition from the box C/D to the Kt-7 kink-turn sequence. A. A plot of the final E_{FRET} values for the box C/D k-turn sequences subsequently modified with the Kt-7 sequence elements in the order of most to least significant, i.e. -1b:-1n=C:G, 3b:3n=A:G, 4b:4n=C:G and L=GAA. B. E_{FRET} (y-axis) has been plotted as a function of Mg^{2+} ion concentration (x-axis) and the curves fitted to the two-state model for the wild-type box C/D (black squares), -1b:-1n=C:G modified box C/D (red circles), -1b:-1n=C:G + 3b:3n=A:G modified box C/D (blue triangles) and -1b:-1n=C:G + 3b:3n=A:G + 4b:4n=C:G + L=GAA modified box C/D (green triangles).

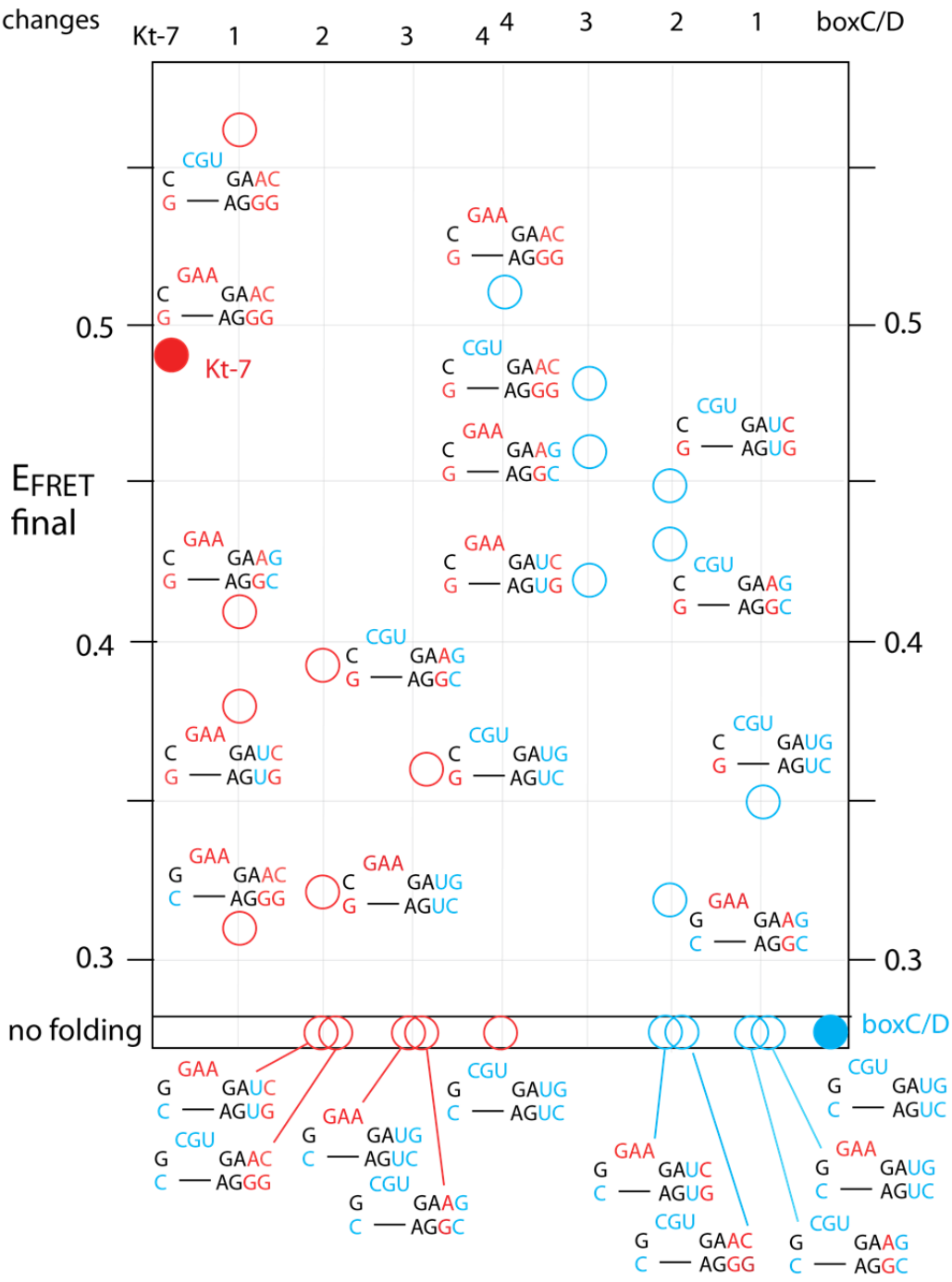


Figure 26: Plot of final E_{FRET} values for all Kt-7 and box C/D variant sequences analysed by FRET for folding properties in Mg^{2+} ions. The final E_{FRET} values for the Kt-7 and box C/D sequences are plotted on the left and right halves of the diagram, respectively. The unmodified Kt-7 and box C/D sequences are shown as red and blue filled circles to the extreme left and right sides of the diagram, respectively. The number of modifications in the Kt-7 sequence increase moving from the left to middle of the diagram. Similarly, the number of modifications in the box C/D sequence increase moving from right to middle of the diagram. Modified Kt-7 and box C/D sequences are represented by red and blue outlined circles, respectively.

The incorporation of box C/D sequence elements into Kt-7 results in the overall reduction of the final E_{FRET} from 0.49 to <0.30 (Fig. 23A), i.e. a well folding k-turn is converted into one with no folding ability in Mg^{2+} ions. The successive incorporation of the box C/D sequence elements, 4b:4n=G:C, 3b:3n=U:U and -1b:-1n=G:C, results in the decrement of the final E_{FRET} value showing their repressive effect on k-turn folding in Mg^{2+} ions and that they work additively to completely prevent folding of the k-turn. The exception is the L=CGU sequence element which results in a slight increase in the final E_{FRET} value suggesting that it has an enhancing effect on k-turn folding in Mg^{2+} ions compared to the L=GAA sequence element. Nevertheless, the greatest effect is exerted by the -1b:-1n=G:C sequence element, the incorporation of which leads to the prevention of k-turn folding in Mg^{2+} ions. Moreover, the incorporation of the -1b:-1n=G:C sequence element along with just one other box C/D sequence element, either 3b:3n=U:U, 4b:4n=G:C or L=CGU, leads to the complete abolishment of k-turn folding in Mg^{2+} ions (Fig. 25). For example, Figure 23B shows that incorporation of the -1b:-1n=G:C sequence element into the 3b:3n=U:U modified Kt-7 k-turn reduces the final E_{FRET} value from 0.38 to <0.30, i.e. prevents k-turn folding in Mg^{2+} ions. Therefore, the magnitude of repressive effect exerted by the -1b:-1n=G:C sequence element is greater than the promotive effect of the 4b:4n=C:G sequence element in the context of the Kt-7 sequence. Thus, overall, the -1b:-1n position exerts the greatest effect on k-turn folding properties in metal ions and therefore is of utmost importance.

Conversely, the incorporation of Kt-7 sequence elements into the box C/D k-turn result in an overall increase of the final E_{FRET} value from below 0.30 to 0.51 (Fig. 24), i.e. a non-folding k-turn is converted into a completely folded k-turn in Mg^{2+} ions. The successive incorporation of the Kt-7 sequence elements, -1b:-1n=C:G, 3b:3n=A:G, 4b:4n=C:G and L=GAA, result in the increment of the final E_{FRET} value confirming their promotive effect on k-turn folding in Mg^{2+} ions and that they work additively to completely fold the k-turn into the kinked conformation. Interestingly, in this case the GAA loop has a promotive effect on k-turn folding in Mg^{2+} ions, as opposed to the promotive effect of the CGU loop in Kt-7 mentioned above. This suggests that the loop

can have both small repressive and promotive effects, and is generally not a significant determinant of the folding characteristics of k-turns in metal ions. Nevertheless, of the sequence elements exchanged, the greatest effect is exerted by the -1b:-1n=C:G sequence element which not only results in the biggest increase in the final E_{FRET} value of box C/D, but also is the sole sequence element which, upon incorporation, can induce folding of the singly modified box C/D k-turn (Fig. 25). Thus, the -1b:-1n position is the most critical sequence determinant for k-turn folding properties in metal ions.

Overall, the -1b:-1n, 3b:3n and 4b:4n positions all influence k-turn folding properties in metal ions. The relatively small and non-definitive effect of the loop on k-turn folding in Mg^{2+} ions renders it to be rather negligible. On the other hand, the -1b:-1n position has the most influence on k-turn folding properties with all the -1b:-1n=C:G sequence element-containing k-turns folding to some extent in Mg^{2+} ions, while -1b:-1n=G:C sequence element-containing k-turns are mostly unable to fold in Mg^{2+} ions. Thus, the -1b:-1n position is the most critical sequence determinant for k-turn folding properties in metal ions. The 3b:3n and 4b:4n positions also influence k-turn folding properties in metal ions although to a lower extent than the -1b:-1n position, hence have intermediate importance in determining folding characteristics of k-turns in metal ions.

	-1b,-1n	Loop	3b,3n	4b,4n	Initial E_{FRET}	Final E_{FRET}	+ L7Ae E_{FRET}
Kt-7	CG	GAA	AG	CG	0.21	0.49	0.63
1 change	GC				0.14	0.31	
	CI				0.21	0.38	
		CGU			0.33	0.56	
		SAA				NF	0.45
		GSA				NF	0.18
		GAS			0.24	0.57	0.67
			UU		0.23	0.38	
				GC	0.14	0.41	
2 changes	GC	CGU				NF	0.64
	GC	AAA				NF	0.62
	GC	CAA				NF	0.60
	GC	SAA				NF	0.26
	GC		UU			NF	0.62
	GC		AC			NF	0.56
	GC		UG		0.36	0.50	0.54
	GC			GC	0.19	0.29	
		CGU	UU		0.28	0.52	0.64
		CGU		GC	0.25	0.39	
			UU	GC	0.22	0.32	
	CI		UU		0.27	0.36	
3 changes	GC	CGU		GC		NF	
	GC		UU	GC		NF	
	GC	CGU	UG		0.28	0.44	0.58
		CGU	UU	GC	0.24	0.36	0.64
4 changes	GC	CGU	UU	GC		NF	0.58
box C/D	GC	CGU	UU	GC		NF	0.63
1 change	CG				0.31	0.35	
		GAA				NF	
			AG			NF	
				CG		NF	
2 changes	CG	GAA			0.29	0.30	
	CG		AG		0.34	0.43	
	CG			CG	0.37	0.45	
		GAA	AG		0.27	0.32	
		GAA		CG		NF	
			AG	CG		NF	
3 changes	CG	GAA	AG		0.26	0.46	0.56
	CG	GAA		CG	0.36	0.42	
	CG		AG	CG	0.40	0.48	
4 changes	CG	GAA	AG	CG	0.32	0.51	0.61

Table 10: Initial and final FRET efficiencies of *H. marismortui* Kt-7 and *A. fulgidus* box C/D kink-turns and their variants with exchanged sequence elements. The initial and final E_{FRET} values on titration with Mg^{2+} ions, for Kt-7 and box C/D k-turns in which the -1b:-

1n, loop, 3b:3n and 4b:4n sequence elements have been exchanged in different combinations. For some of these constructs, 1μM L7Ae was added after the final addition of Mg^{2+} ions and the value of E_{FRET} is tabulated. NF = no folding. I = inosine. S = spacer (nucleotide with no nucleobase).

5.3 Folding Analysis of the Kt-7 Kink-Turn and its Variants by Comparative Gel Electrophoresis

The folding properties in Mg^{2+} ions of unmodified Kt-7 and its variants with either the -1b:-1n=G:C or L=CGU sequence element or both, were analysed by the alternative method of polyacrylamide gel electrophoresis. The electrophoretic mobility of the RNA in the gel depends on the conformation it adopts (see Materials and Methods). The more bent the k-turn i.e. the more completely folded the k-turn, the greater the electrophoretic retardation of the k-turn species in the gel [67]. The constructs used in this method of study comprise of 65bp DNA/RNA/DNA duplexes which contain a central core of k-turn-containing RNA. The complete sequences of these duplexes are presented in Materials and Methods. The samples were electrophoresed in a 13% polyacrylamide gel in non-denaturing conditions in the presence of either 2mM EDTA or 2mM $MgCl_2$. The gel electrophoresis analysis of Kt-7 and its variants is shown in Figure 26. To provide a reference, a mixture of two equivalent duplexes containing the A_3 and A_7 bulges in place of the k-turns were electrophoresed alongside the Kt-7 samples. The A_7 bulge kinks the axis of the helix in a manner similar to that of a well folded k-turn, while the A_3 bulge results in a smaller kink that is similar to an unfolded k-turn.

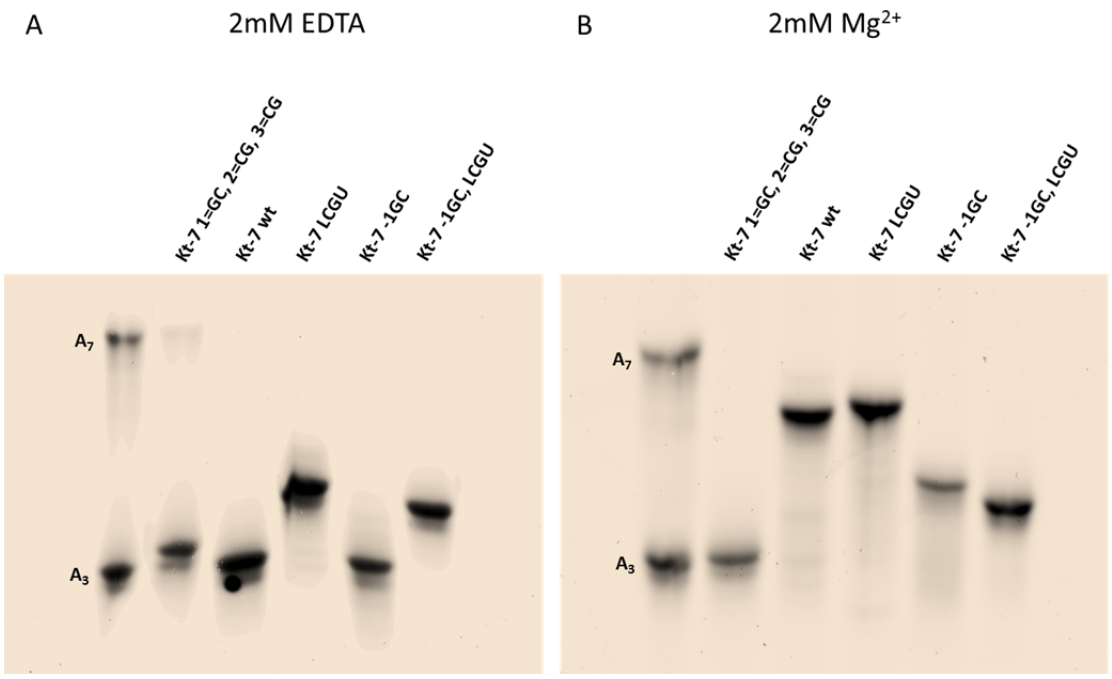


Figure 27: Analysis of the Mg^{2+} ion-dependent folding of *H. marismortui* Kt-7 and its variants by gel electrophoresis. The 65bp DNA/RNA/DNA duplexes with centrally located k-turn-containing RNA regions were electrophoresed in a 13% polyacrylamide gel in the presence of A. 2mM EDTA or B. 2mM Mg^{2+} ions. For reference, two equivalent duplexes, in which the k-turns were replaced by the A_3 or A_7 bulges, were mixed and electrophoresed in lane 1. Lane 2 contains a Kt-7 variant in which the three G:A base pairs have been changed to Watson-Crick G:C base pairs by the substitutions of A1nC, A2bC and A3bC. Lanes 3 to 6 contain Kt-7 and its variants modified with sequence elements taken from the *Af* boxC/D k-turn. Lane 3 is the wild-type Kt-7; lane 4 is the L=CGU modified Kt-7; lane 5 is the -1b:-1n=G:C modified Kt-7; and lane 6 is the -1b:-1n=G:C + L=CGU modified Kt-7.

A Kt-7 variant with Watson-Crick base pairs in positions 1, 2 and 3 (A1nC, A2bC and A3bC, i.e. this is a simple 3-nucleotide bulge) co-migrates with the A_3 bulge in gels electrophoresed in both EDTA (Fig. 26A, lane 2) and Mg^{2+} ions (Fig. 26B, lane 2). This suggests that this Kt-7 variant is unable to fold into the kinked conformation even in the presence of Mg^{2+} ions and behaves like a simple three nucleotide bulge-containing RNA molecule. Certainly, this is consistent with the requirement of *trans* G(sugar edge):A(Hoogsteen edge) base pairs in positions 1 and 2 of the NC helix to determine a standard k-turn structural motif with the ability to undergo metal ion-induced folding into the

kinked conformation[37, 54]. Also, the requirement of a non-Watson-Crick base pair in the 3b:3n position to induce folding in the presence of Mg^{2+} ions[66] has previously been explored. Thus, the Watson-Crick base pairs in positions 1, 2 and 3 render this Kt-7 variant unable to fold in Mg^{2+} ions.

The wild-type Kt-7 k-turn co-migrates with the A_3 bulge and the A_7 bulge in the absence of metal ions (Fig. 26A, lane 3) and presence of metal ions (Fig. 26B, lane 3), respectively. This implies that the unmodified Kt-7 k-turn is unable to fold in the absence of Mg^{2+} ions but does fold well into the kinked conformation in the presence of Mg^{2+} ions, consistent with the FRET results. Modifying the loop of Kt-7 to CGU results in greater electrophoretic retardation than the unmodified Kt-7 in the presence of both EDTA (Fig. 26A, lane 4) and Mg^{2+} ions (Fig. 26B, lane 4). However, the magnitude of retardation is much greater in the presence of EDTA than in the presence of Mg^{2+} ions in which there is only slightly more retardation. This implies that the L=CGU modified Kt-7 has a more bent structure than the unmodified Kt-7 in the unfolded state in the absence of Mg^{2+} ions, consistent with the higher initial E_{FRET} value of the L=CGU modified Kt-7 compared to the unmodified k-turn (Fig. 16). Furthermore, the small increase in the final E_{FRET} value of the L=CGU modified Kt-7 compared to the unmodified Kt-7 (Fig. 16) supports the slightly greater electrophoretic retardation of the L=CGU modified Kt-7 in the presence of Mg^{2+} ions. Again, this implies the small positive effect of the CGU loop on k-turn folding in Mg^{2+} ions although to a very small extent.

The folding of the -1b:-1n=G:C modified Kt-7 in the presence of Mg^{2+} ions is severely diminished, with a migration intermediate between the unmodified Kt-7 and the A_3 bulge (Fig. 26B, lane 5). This implies that the -1b:-1n=G:C modified Kt-7 does fold to some extent in Mg^{2+} ions but not completely like the wild-type Kt-7, consistent with the FRET results (Fig. 13). Similarly, the -1b:-1n=G:C + L=CGU modified Kt-7 has lower electrophoretic retardation compared to both the wild-type Kt-7 and the -1b:-1n=G:C modified Kt-7, in the presence of Mg^{2+} ions (Fig. 26B, lane 6). This again implies the impairment of k-turn folding in Mg^{2+} ions, consistent with the FRET results. On the other hand, the -1b:-1n=G:C + L=CGU modified Kt-7 has greater electrophoretic retardation than the wild-type and the -1b:-1n=G:C modified

Kt-7 in the presence of EDTA (Fig. 26A, lane 6). This construct migrates with the L=CGU modified Kt-7 suggesting a similarly more bent structure of the k-turn in the unfolded state in the absence of Mg^{2+} ions. Again, this is due to the presence of CGU loop in the k-turn sequence which, in the unfolded state, produces a greater kink in the axis of the RNA molecule than the GAA loop.

Overall, it can be stated that wild-type Kt-7 and its variants are unfolded in the absence of Mg^{2+} ions. However, the CGU loop-containing k-turns are more bent in the unfolded state than the GAA loop-containing k-turns. Nevertheless, -1b:-1n=C:G sequence element-containing k-turns undergo complete Mg^{2+} ion-induced folding into the kinked conformation irrespective of the identity of the loop. This implies that the -1b:-1n=C:G sequence element promotes Mg^{2+} ion-induced folding of the k-turn and that its effect is much greater in magnitude than that of the loop. On the other hand, the -1b:-1n=G:C sequence element-containing k-turns undergo limited Mg^{2+} ion-induced folding into the k-turn conformation. This implies the repressive effect of the -1b:-1n=G:C sequence element on k-turn folding properties in Mg^{2+} ions. Thus, on the whole, in opposition to the negligible loop, the -1b:-1n position is a key sequence determinant for k-turn folding properties in metal ions.

5.4 The Effect of the N^6-methyladenine Modification on Kt-7 Folding Properties Analysed by FRET

The adenine nucleobases in positions 1n, 2b, 3b, as well as L3, in the Kt-7 k-turn sequence were individually substituted by N^6-methyladenine (m6A). The modified k-turns were then analysed for their folding properties in Mg^{2+} ions using FRET. The E_{FRET} data, as a function of Mg^{2+} ion concentration, for the modified Kt-7 k-turns is plotted in Figure 27, alongside the E_{FRET} data for the unmodified Kt-7 k-turn. After the final addition of Mg^{2+} ions, 1μM L7Ae protein was added to the k-turn RNA to observe its protein-induced folding ability. The E_{FRET} value obtained after the addition of L7Ae to each construct,

as well as the initial and final E_{FRET} values observed with the Mg^{2+} ion titrations, is listed in Table 2.

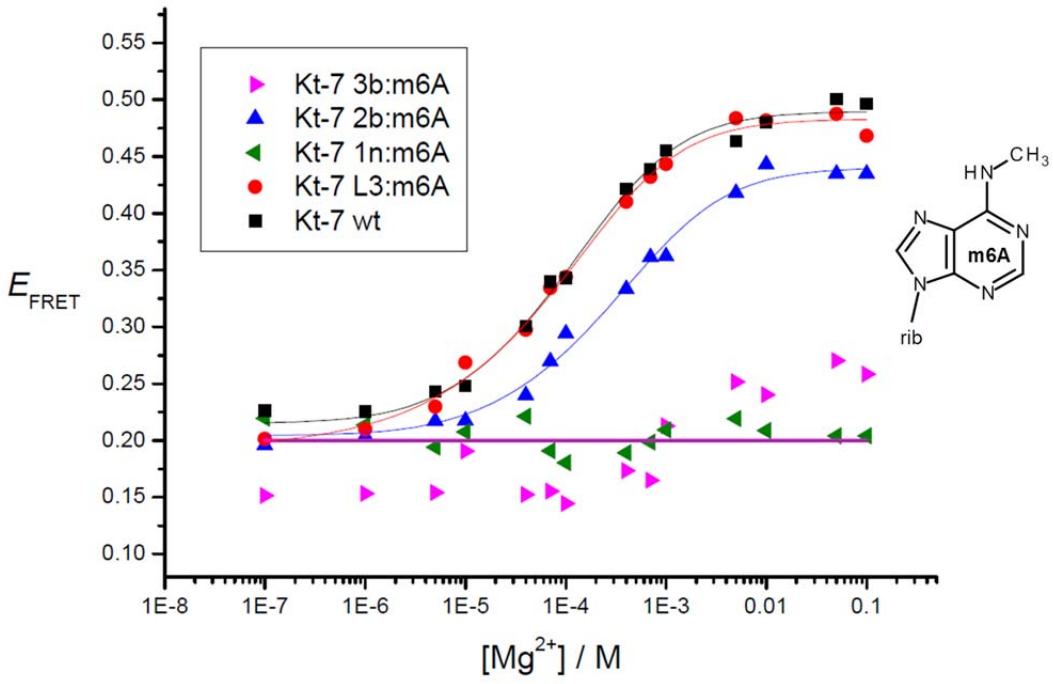


Figure 28: Mg^{2+} ion-dependent folding analysis of wild-type *H. marismortui* Kt-7 and its N^6-methyladenine modified variants by FRET. E_{FRET} (y-axis) has been plotted as a function of Mg^{2+} ion concentration (x-axis) and the data for the wild-type Kt-7 (black squares), L3:m6A modified Kt-7 (red circles), 1n:m6A modified Kt-7 (green triangles), 2b:m6A modified Kt-7 (blue triangles) and 3b:m6A modified Kt-7 (pink triangles) are fitted to the two-state model. The structure of m6A is displayed on the right.

	Initial E_{FRET}	Final E_{FRET}	+ L7Ae E_{FRET}
Kt-7 wt	0.21	0.49	0.63
Kt-7 1n:m6A	NF		0.39
Kt-7 2b:m6A	0.20	0.44	0.62
Kt-7 3b:m6A	NF		0.67
Kt-7 L3:m6A	0.19	0.48	0.62

Table 11: FRET analysis of Mg^{2+} ion and L7Ae protein-dependent folding of the wild-type *H. marismortui* Kt-7 kink-turn and its N^6-methyladenine modified variants. The initial and final E_{FRET} values obtained on titration with Mg^{2+} ions, as well as the E_{FRET} value

obtained after the addition of 1μM L7Ae, are tabulated for the unmodified Kt-7 k-turn and for the various Kt-7 k-turns modified with the N^6-methyladenine substitution at positions 1n, 2b, 3b and L3. NF = no folding.

The N^6-methylation of the adenine in position L3 of Kt-7 leads to no change in the k-turn folding properties relative to the unmodified Kt-7 k-turn. The L3:m6A modified and unmodified constructs fold similarly well in Mg^{2+} ions with final E_{FRET} values of 0.49 and 0.48, respectively (Fig. 26). This suggests that the m6A modification of the L3 nucleobase does not perturb the k-turn's folding ability which is consistent with the non-involvement of the L3 nucleobase in any k-turn stabilizing interactions (Fig. 6) making this position tolerant of any modifications. This is further confirmed by the complete removal of the nucleobase from the L3 position of the Kt-7 k-turn sequence which, also, still allows the construct to undergo Mg^{2+} ion-induced folding into the kinked conformation similar to the unmodified Kt-7, as observed by FRET (Appendix Fig. 1C). Furthermore, the protein-induced folding ability is also not perturbed by the L3:m6A modification as revealed by the E_{FRET} value of 0.62 obtained upon L7Ae binding, comparable to the 0.63 E_{FRET} value of the unmodified Kt-7 (Table 2). Thus, overall, the N^6-methylation of the L3:A nucleobase does not affect the folding properties of the k-turn.

On the other hand, the N^6-methylation of the adenine in the 1n position of Kt-7 results in complete abolishment of k-turn folding ability in Mg^{2+} ions as revealed by the reduction of the final E_{FRET} value from 0.49 to <0.30 (Fig. 26). Furthermore, the L7Ae protein-induced folding ability is also impaired as depicted by the decrease in the E_{FRET} value, obtained upon addition of 1μM L7Ae, from 0.63 to 0.39 (Table 2). This clearly suggests the strongly perturbing effect of the m6A modification of the A1n nucleobase which not only converts a well folded k-turn into one with no folding ability in Mg^{2+} ions, but also impairs its protein-induced folding ability. Thus, this 1n modification has a marked effect on the k-turn folding properties.

Similarly, the N^6-methylation of the adenine in the 3b position of Kt-7 also results in complete abolishment of k-turn folding ability in Mg^{2+} ions as revealed by the reduction of the final E_{FRET} value from 0.49 to <0.30 (Fig. 26), converting a well folded k-turn into one with no folding ability in Mg^{2+}

ions. However, in this case, the protein-induced folding ability is not affected as the 3b:m6A modified construct folds just as well with an E_{FRET} value of 0.67 upon L7Ae binding, comparable to the 0.63 E_{FRET} value of the unmodified construct (Table 2). Thus, the 3b:m6A modification strongly affects the Mg^{2+} ion-induced folding ability of the k-turn but not the protein-induced folding ability. Therefore, this 3b modification has implications for k-turn folding properties in metal ions alone.

The N^6-methylation of the adenine in the 2b position of Kt-7 has a rather smaller effect on Mg^{2+} ion-induced folding ability. The final E_{FRET} value is reduced from 0.49 to 0.44 (Fig. 26), suggesting only slight impairment of k-turn folding in Mg^{2+} ions. The protein-induced folding ability is also not affected with an E_{FRET} value of 0.62 depicted upon L7Ae binding, comparable to the 0.63 E_{FRET} value of the unmodified construct (Table 2). Thus, the m6A modification of the A2b nucleobase only slightly affects the k-turn folding properties in metal ions, while the protein-induced folding ability remains unaffected. This implies the little significance of this modification in the 2b position of the k-turn sequence on determination of k-turn folding properties.

Overall, the m6A modification has a variable effect on k-turn folding properties dependent on its location in the k-turn sequence. In the L3 and 2b positions this modification is tolerated resulting in little or no influence on k-turn folding properties in Mg^{2+} ions. In the 1n and 3b positions, however, the m6A modification results in complete prevention of k-turn folding in Mg^{2+} ions. Thus, in these two positions, the m6A modification exerts a strong detrimental effect on k-turn folding properties in metal ions. Nevertheless, it is in the 1n position that the m6A modification exerts its maximal effect on k-turn folding properties, with neither metal ion-dependent or protein-induced folding occurring. By contrast, the protein-induced folding ability remains unaffected for the m6A modification in the 3b position, as well as in the L3 and 2b positions. Thus, overall, this variable effect of the m6A modification renders the 1n position of the k-turn sequence to be the most influential for general k-turn folding properties upon N^6-adenine methylation, while the 3b position has implications for k-turn folding properties in metal ions alone, as opposed to the rather non-significant L3 and 2b positions.

6. Discussion

The ability of the ribosomal *Hm*Kt-7 k-turn to fold in metal ions, and the inability of the *Af*boxC/D snoRNA k-turn to fold in metal ions, has revealed the existence of sequence elements that influence the folding properties of k-turns. The 3b:3n has previously been identified as a critical determinant of metal ion-induced k-turn folding. However, the U3b:U3n sequence element of the *Af*boxC/D k-turn allowed moderate metal ion-induced folding of the 3b:3n=U:U modified *Hm*Kt-7. This indicated the existence of sequence elements, other than 3b:3n, which also affect the folding properties of k-turns in metal ions. The comparison of the *Hm*Kt-7 and *Af*boxC/D k-turn sequences revealed three other sequence elements that differed between the two k-turns, which could potentially affect their folding properties. These were the -1b:-1n and 4b:4n positions, as well as the loop. The exchange of these sequence elements between the *Hm*Kt-7 and *Af*boxC/D k-turns has revealed their relative importance for the determination of standard k-turn folding properties in response to the addition of metal ions. While the loop has a negligible effect, the -1b:-1n, 3b:3n and 4b:4n positions all affect the k-turn folding properties in an approximately additive manner. However, the largest effect is exerted by the -1b:-1n position, which is the major contributor amongst all positions studied, towards k-turn folding properties in Mg^{2+} ions. Hence, between the two k-turns that are compared, the -1b:-1n position is the most critical sequence determinant for k-turn folding properties in metal ions, while the 3b:3n and 4b:4n positions have intermediate importance for influencing k-turn folding. Note that this conclusion is based on the study of only two alternative sequence elements in each of the positions.

Interestingly, *Hm*Kt-7 has selected sequence elements, C-1b:G-1n, A3b:G3n and C4b:G4n, which induce the folding of the k-turn in metal ions. The C-1b:G-1n and C4b:G4n sequence elements, as well as the G3n sequence element which coordinates Mg^{2+} ions to stabilise the k-turn conformation [66], are strongly conserved in all bacterial Kt-7 sequences and more generally in ribosomal k-turns [112]. This may reflect their biological requirement for the formation of the k-turn structure in metal ions alone, for

the unhindered assembly of the ribosome. Indeed within the ribosome *Hm*Kt-7 binds the L24 protein [37] but surely this is not necessary to fold the k-turn. Rather, the L24 protein probably binds to the folded k-turn structure to accomplish the subsequent assembly of the ribosome.

Similarly to ribosomes, a number of riboswitches also recruit k-turns as structural elements. For these there is no known binding protein, so they must be able to fold in metal ions alone. Therefore, these will require sequence elements that allow folding into the kinked conformation in metal ions alone, for the unhindered assembly of the riboswitch. For example, the SAM-I riboswitch k-turn has strongly conserved the A3b:G3n and C-1b:G-1n sequence elements [66, 112] which promote the metal ion-induced folding of the k-turn. Hence, the SAM-I riboswitch k-turn has a strong tendency to undergo folding into the kinked conformation in metal ions alone. This clearly is a requirement of the SAM-I riboswitch k-turn and is consistent with the absence of any known protein that binds to this riboswitch. Thus, in general, riboswitch k-turns can fold into the kinked conformation, in the absence of proteins, in metal ions alone by the selection of the appropriate sequence elements.

The *Af*boxC/D k-turn, on the other hand, has selected sequence elements, G-1b:C-1n, U3b:U3n and G4b:C4n, which do not facilitate k-turn folding in metal ions. The U3b:U3n and G4b:C4n sequence elements are strongly conserved in all archaeal and human box C/D k-turns [55, 112]. However, the G-1b:C-1n sequence element is less conserved with the C:G base pair predominantly found in the -1b:-1n position of the box C/D k-turn sequence. Nevertheless, the combination of the C-1b:G-1n sequence element with the U3b:U3n and G4b:C4n sequence elements still results in the very poor folding of the k-turn in metal ions (Table 1). Thus, the *Af*boxC/D k-turn is unable to fold in metal ions alone and requires the binding of the L7Ae-type protein to assist folding into the kinked conformation. This again may have biological significance such that free formation of the k-turn structure may not be necessary, but rather folding into the kinked conformation shall take place upon binding of the L7Ae-type protein, the recruitment of which is actually

the first crucial step in the assembly of the biologically active box C/D snoRNP complex [83].

While the effect of only two base pairs in each of the -1b:-1n, 3b:3n and 4b:4n positions has been studied (i.e. those found in *Hm*Kt-7 and *Af*boxC/D), surely there will exist other base pairs in these various positions. For example, in the standard k-turns, *H. marismortui* ribosomal Kt-46 has a G:U base pair in the 4b:4n position, *H. sapiens* U4 Kt has a G:C base pair in the 3b:3n position, and *B. subtilis* T-box riboswitch k-turn has a U:A base pair in the -1b:-1n position. It would be interesting to investigate how these other base pairs fare towards k-turn folding properties in metal ions. The effect of all sixteen possible base pair compositions in the 3b:3n position of *Hm*Kt-7 on folding properties in Mg^{2+} ions has previously been extensively studied [66] (Appendix Fig. 2). From our study, the well folding and intermediately folding, 3b:3n=A:G and 3b:3n=U:U sequence element-containing Kt-7 k-turns, respectively, are consistent with this previous research. A similar extensive study of the -1b:-1n and 4b:4n positions could provide a valuable insight into the effect of the various base pairs in these two positions on the k-turn's folding properties in metal ions, thus providing a stronger predictive tool for standard k-turn folding in metal ions.

By far, the -1b:-1n position has been identified as the most critical sequence determinant for k-turn folding properties in metal ions. The reason for this criticality lies in the role of the -1n position to form key k-turn stabilising A-minor interactions with the position A2b. Evidently, the preference of C-1b:G-1n, that facilitates metal ion-induced folding, over the G-1b:C-1n sequence element (that suppresses it), suggests the ability of C-1b:G-1n to form a stronger A-minor interaction with the position A2b than the G-1b:C-1n sequence element. In particular, the impairment of k-turn folding in Mg^{2+} ions upon removal of the exocyclic amine group from the G-1n nucleobase suggests the involvement of this exocyclic amine group in the formation of the stronger A-minor interaction. Figure 29 shows that the exocyclic amine of G-1n forms an additional, k-turn stabilising, cross-strand hydrogen bond with A2b in the core of the k-turn. The N2 atom of the G-1n nucleobase forms a hydrogen bond with the O2' atom of A2b, in addition to the hydrogen bond

formed between O2' atom of G-1n and the A2b nucleobase, which is also formed by the C nucleotide in the -1n position when there is a G:C base pair in the -1b:-1n position. Thus, upon removal of the exocyclic amine group from the G-1n nucleobase, the relative hydrogen bond is disrupted weakening the A-minor interaction to a similar extent as to the G-1b:C-1n sequence element, hence the k-turn structure is destabilised. Therefore, the preference of the C:G base pair in the -1b:-1n position can be attributed to the ability of the G-1n nucleobase to form an additional cross-strand hydrogen bond with A2b and, therefore, a stronger k-turn stabilising A-minor interaction.

Figure 29: The A-minor interaction of G-1n with A2b in *H. marismortui* Kt-7. The two cross-strand hydrogen bonds are coloured red. One hydrogen bond is donated by the O2' atom of G-1n ribose to the N3 atom of the A2b nucleobase. The second hydrogen bond is formed between the N2 atom of G-1n nucleobase and the O2' atom of the A2b ribose.

While the base pair composition of the various positions in the k-turn sequence can affect k-turn folding properties in metal ions, the N^6-methylation of the A nucleobase has also been shown to affect the Mg^{2+} ion-induced folding of the standard Kt-7 k-turn. While, N^6-adenine methylation in positions L3 and 2b has none or very little effect on Mg^{2+} ion-induced folding of the k-turn, in the 1n and 3b positions it completely prevents folding of the k-turn in Mg^{2+} ions via the disruption of the *trans* G(sugar edge):A(Hoogsteen edge) base pair required to form the k-turn structure [55]. Thus, N^6-adenine methylation at 1n and 3b plays an important role in influencing standard k-turn folding properties in metal ions should it be naturally occurring.

The consensus sequence for the METTL3-METTL14 methyltransferase enzyme that generates N^6-methyladenine in eukaryotes can be simplified to GAC [98], where the central A is methylated. A subset of box C/D k-turns, that have the G-1b:C-1n sequence element (e.g. *Af*boxC/D studied here), have this GAC sequence motif in the [2n, 1n, -1n] region. And, interestingly, several such human box C/D k-turns have naturally been found to be N^6-methylated at the A1n position [55]. Surely, the G-1b:C-1n sequence element is itself sufficient to impair k-turn folding in metal ions suggesting an alternative role for the 1n:m6A modification. Certainly, the N^6-adenine methylation of the 1n position has been found to prevent binding of the 15.5k protein to the human box C/D k-turn and impair the folding of the k-turn [55], consistent with this research. Therefore, although the G-1b:C-1n sequence element itself does not affect the protein-induced folding of the k-turn but rather only impairs the metal ion-induced folding of the k-turn, it does predispose the k-turn to be methylated at the 1n position which prevents the L7Ae-type protein from binding and therefore prevents the protein-induced folding of the k-turn. Thus, the G-1b:C-1n sequence element directly affects the metal ion-induced folding of the k-turn and indirectly affects the protein-induced folding of the k-turn via the recruitment of the 1n:m6A modification. Surely, in general, box C/D k-turns are unable to fold in metal ions alone due to the selection of repressive sequence elements in the 3b:3n and 4b:4n positions, irrespective of the base pair in the -1b:-1n position. Nevertheless, the recruitment of the G-1b:C-1n sequence element allows the further regulation of the protein-induced folding of the k-turn. Overall, this reveals some potentially significant regulatory pathways that the box C/D k-turn adopts via the selection of sequence elements and subsequent modifications to control the formation of the k-turn structure and the subsequent assembly of the functionally active box C/D complex.

Like the box C/D k-turn, the U4 snRNA k-turn also recruits a 3b:3n sequence element that inhibits the metal ion-induced folding of the k-turn. However, while box C/D recruits the intermediate folding U3b:U3n sequence element which together with the repressive G4b:C4n sequence element strongly impairs the metal ion-induced folding of the k-turn, the U4 snRNA k-turn

strongly conserves the G3b:C3n sequence element [66] which itself completely prevents the metal ion-induced folding of the k-turn. On the contrary, U4 snRNA k-turns have conserved the C-1b:G-1n sequence element [112] as opposed to the G-1b:C-1n sequence element which seems largely restricted to box C/D k-turns. Despite this, the U4 snRNA k-turn is unable to fold in metal ions alone and, like box C/D, requires the assistance of the L7Ae-type protein to fold. This again will be in requirement to its biological duty which, like box C/D, does not require the free formation of the k-turn structure in metal ions alone, but rather the generation of a ribonucleoprotein complex, the spliceosome. However, in general, unlike selective box C/D k-turns, the U4 snRNA k-turn is unable to regulate k-turn folding upon protein binding via N^6-adenine methylation at the 1n position, due to the lack of the G-1b:C-1n sequence element. Hence, U4 snRNA has avoided this regulatory pathway in the biogenesis of the spliceosome.

In summary, many sequence elements in the k-turn contribute towards its folding properties. The -1b:-1n, 3b:3n and 4b:4n sequence elements affect k-turn folding properties in an approximately additive manner with the ultimate effect exerted by the -1b:-1n position. Different classes of k-turns select different sequence elements that are best adapted to their specific biological requirement. Those that are required to freely fold in the presence of metal ions alone select different sequence elements from those that are involved in the biogenesis of ribonucleoprotein complexes. In general, ribosomal and riboswitch k-turns select promotive sequence elements that facilitate folding in metal ions, while box C/D and U4 snRNA k-turns select sequence elements that repress their folding ability in metal ions rendering them foldable only upon binding of the appropriate protein. Nevertheless, k-turns can also adopt modifications, such as the N^6-adenine methylation, to influence k-turn folding properties. The recruitment of the 1n:m6A modification by box C/D, for example, prevents the protein-induced folding of the k-turn, thus potentially providing a regulatory pathway for the assembly of the biologically active box C/D snoRNP complex. Overall, this reveals the widening ability of k-turns to control their folding properties by the selection of

specific sequence elements pre-transcriptionally and by the recruitment of modifications post-transcriptionally.

7. List of References

1. Crick, F., *Central dogma of molecular biology.* Nature, 1970. **227**(5258): p. 561-3.
2. Brenner, S., F. Jacob, and M. Meselson, *An unstable intermediate carrying information from genes to ribosomes for protein synthesis.* Nature, 1961. **190**: p. 576-581.
3. Eddy, S.R., *Non-coding RNA genes and the modern RNA world.* Nature reviews. Genetics, 2001. **2**(12): p. 919-29.
4. Crick, F.H., *On protein synthesis.* Symposia of the Society for Experimental Biology, 1958. **12**: p. 138-63.
5. Hoagland, M.B., et al., *A soluble ribonucleic acid intermediate in protein synthesis.* The Journal of biological chemistry, 1958. **231**(1): p. 241-57.
6. Lilley, D.M., *Structure, folding and mechanisms of ribozymes.* Current opinion in structural biology, 2005. **15**(3): p. 313-23.
7. Serganov, A. and D.J. Patel, *Ribozymes, riboswitches and beyond: regulation of gene expression without proteins.* Nature reviews. Genetics, 2007. **8**(10): p. 776-90.
8. Serganov, A. and E. Nudler, *A decade of riboswitches.* Cell, 2013. **152**(1-2): p. 17-24.
9. Breaker, R.R., *Prospects for riboswitch discovery and analysis.* Molecular cell, 2011. **43**(6): p. 867-79.
10. Lilley, D.M., *The Guanidine Riboswitch-A Poor Orphan No Longer.* Cell chemical biology, 2017. **24**(2): p. 130-131.
11. Liu, N., et al., *Probing N6-methyladenosine RNA modification status at single nucleotide resolution in mRNA and long noncoding RNA.* RNA, 2013. **19**(12): p. 1848-56.
12. Zhou, K.I., et al., *N(6)-Methyladenosine Modification in a Long Noncoding RNA Hairpin Predisposes Its Conformation to Protein Binding.* Journal of molecular biology, 2016. **428**(5 Pt A): p. 822-833.
13. Busch, H., et al., *SnRNAs, SnRNPs, and RNA processing.* Annual review of biochemistry, 1982. **51**: p. 617-54.

14. Tarn, W.Y. and J.A. Steitz, *Highly diverged U4 and U6 small nuclear RNAs required for splicing rare AT-AC introns.* Science, 1996. **273**(5283): p. 1824-32.
15. He, L. and G.J. Hannon, *MicroRNAs: small RNAs with a big role in gene regulation.* Nature reviews. Genetics, 2004. **5**(7): p. 522-31.
16. Elbashir, S.M., W. Lendeckel, and T. Tuschl, *RNA interference is mediated by 21- and 22-nucleotide RNAs.* Genes & development, 2001. **15**(2): p. 188-200.
17. Carthew, R.W., *Gene silencing by double-stranded RNA.* Current opinion in cell biology, 2001. **13**(2): p. 244-8.
18. Sharp, P.A., *RNA interference--2001.* Genes & development, 2001. **15**(5): p. 485-90.
19. Maxwell, E.S. and M.J. Fournier, *The small nucleolar RNAs.* Annual review of biochemistry, 1995. **64**: p. 897-934.
20. Balakin, A.G., L. Smith, and M.J. Fournier, *The RNA world of the nucleolus: two major families of small RNAs defined by different box elements with related functions.* Cell, 1996. **86**(5): p. 823-34.
21. Kiss-Laszlo, Z., et al., *Site-specific ribose methylation of preribosomal RNA: a novel function for small nucleolar RNAs.* Cell, 1996. **85**(7): p. 1077-88.
22. Ganot, P., M.L. Bortolin, and T. Kiss, *Site-specific pseudouridine formation in preribosomal RNA is guided by small nucleolar RNAs.* Cell, 1997. **89**(5): p. 799-809.
23. Ni, J., A.L. Tien, and M.J. Fournier, *Small nucleolar RNAs direct site-specific synthesis of pseudouridine in ribosomal RNA.* Cell, 1997. **89**(4): p. 565-73.
24. Esquela-Kerscher, A. and F.J. Slack, *Oncomirs - microRNAs with a role in cancer.* Nature reviews. Cancer, 2006. **6**(4): p. 259-69.
25. Mitra, S.A., A.P. Mitra, and T.J. Triche, *A central role for long non-coding RNA in cancer.* Frontiers in genetics, 2012. **3**: p. 17.
26. Papait, R., et al., *Long noncoding RNA: a new player of heart failure?* Journal of cardiovascular translational research, 2013. **6**(6): p. 876-83.
27. Latronico, M.V. and G. Condorelli, *MicroRNAs and cardiac pathology.* Nature reviews. Cardiology, 2009. **6**(6): p. 419-29.

28. Holley, R.W., et al., *Structure of a Ribonucleic Acid.* Science, 1965. **147**(3664): p. 1462-5.
29. Rich, A. and D.R. Davies, *A NEW TWO STRANDED HELICAL STRUCTURE: POLYADENYLIC ACID AND POLYURIDYLIC ACID.* Journal of the American Chemical Society, 1956. **78**(14): p. 3548-3549.
30. Svoboda, P. and A. Di Cara, *Hairpin RNA: a secondary structure of primary importance.* Cellular and molecular life sciences : CMLS, 2006. **63**(7-8): p. 901-8.
31. Rumnieks, J. and K. Tars, *Crystal structure of the bacteriophage Qbeta coat protein in complex with the RNA operator of the replicase gene.* Journal of molecular biology, 2014. **426**(5): p. 1039-49.
32. Staple, D.W. and S.E. Butcher, *Pseudoknots: RNA structures with diverse functions.* PLoS biology, 2005. **3**(6): p. e213.
33. Rastogi, T., et al., *A long-range pseudoknot is required for activity of the Neurospora VS ribozyme.* The EMBO journal, 1996. **15**(11): p. 2820-5.
34. Adams, P.L., et al., *Crystal structure of a group I intron splicing intermediate.* RNA, 2004. **10**(12): p. 1867-87.
35. Plant, E.P., et al., *The 9-angstrom solution: How mRNA pseudoknots promote efficient programmed -1 ribosomal frameshifting.* RNA, 2003. **9**(2): p. 168-174.
36. Huang, L. and D.M. Lilley, *The Kink Turn, a Key Architectural Element in RNA Structure.* Journal of molecular biology, 2016. **428**(5 Pt A): p. 790-801.
37. Lilley, D.M., *The K-turn motif in riboswitches and other RNA species.* Biochimica et biophysica acta, 2014. **1839**(10): p. 995-1004.
38. Klein, D.J., et al., *The kink-turn: a new RNA secondary structure motif.* The EMBO journal, 2001. **20**(15): p. 4214-21.
39. Montange, R.K. and R.T. Batey, *Structure of the S-adenosylmethionine riboswitch regulatory mRNA element.* Nature, 2006. **441**(7097): p. 1172-5.

40. Zhang, J. and A.R. Ferre-D'Amare, *Co-crystal structure of a T-box riboswitch stem I domain in complex with its cognate tRNA.* Nature, 2013. **500**(7462): p. 363-6.
41. Smith, K.D., et al., *Structural basis of differential ligand recognition by two classes of bis-(3'-5')-cyclic dimeric guanosine monophosphate-binding riboswitches.* Proceedings of the National Academy of Sciences of the United States of America, 2011. **108**(19): p. 7757-62.
42. Peselis, A. and A. Serganov, *Structural insights into ligand binding and gene expression control by an adenosylcobalamin riboswitch.* Nature structural & molecular biology, 2012. **19**(11): p. 1182-4.
43. Hamma, T. and A.R. Ferre-D'Amare, *Structure of protein L7Ae bound to a K-turn derived from an archaeal box H/ACA sRNA at 1.8 A resolution.* Structure, 2004. **12**(5): p. 893-903.
44. Moore, T., et al., *Molecular basis of box C/D RNA-protein interactions; cocrystal structure of archaeal L7Ae and a box C/D RNA.* Structure, 2004. **12**(5): p. 807-18.
45. Szewczak, L.B., et al., *Molecular basis for RNA kink-turn recognition by the h15.5K small RNP protein.* RNA, 2005. **11**(9): p. 1407-19.
46. Youssef, O.A., R.M. Terns, and M.P. Terns, *Dynamic interactions within sub-complexes of the H/ACA pseudouridylation guide RNP.* Nucleic acids research, 2007. **35**(18): p. 6196-206.
47. Vidovic, I., et al., *Crystal structure of the spliceosomal 15.5kD protein bound to a U4 snRNA fragment.* Molecular cell, 2000. **6**(6): p. 1331-42.
48. Wozniak, A.K., et al., *Detecting protein-induced folding of the U4 snRNA kink-turn by single-molecule multiparameter FRET measurements.* RNA, 2005. **11**(10): p. 1545-54.
49. White, S.A., et al., *Internal loop mutations in the ribosomal protein L30 binding site of the yeast L30 RNA transcript.* RNA, 2004. **10**(3): p. 369-77.
50. Reiter, N.J., et al., *Structure of a bacterial ribonuclease P holoenzyme in complex with tRNA.* Nature, 2010. **468**(7325): p. 784-9.

51. Liu, J. and D.M. Lilley, *The role of specific 2'-hydroxyl groups in the stabilization of the folded conformation of kink-turn RNA.* RNA, 2007. **13**(2): p. 200-10.
52. Schroeder, K.T. and D.M. Lilley, *Ion-induced folding of a kink turn that departs from the conventional sequence.* Nucleic acids research, 2009. **37**(21): p. 7281-9.
53. Wang, J., et al., *The k-junction motif in RNA structure.* Nucleic acids research, 2014. **42**(8): p. 5322-31.
54. Turner, B. and D.M. Lilley, *The importance of G.A hydrogen bonding in the metal ion- and protein-induced folding of a kink turn RNA.* Journal of molecular biology, 2008. **381**(2): p. 431-42.
55. Huang, L., et al., *Control of box C/D snoRNP assembly by N6-methylation of adenine.* EMBO reports, 2017. **18**(9): p. 1631-1645.
56. Daldrop, P. and D.M. Lilley, *The plasticity of a structural motif in RNA: structural polymorphism of a kink turn as a function of its environment.* RNA, 2013. **19**(3): p. 357-64.
57. Nissen, P., et al., *RNA tertiary interactions in the large ribosomal subunit: the A-minor motif.* Proceedings of the National Academy of Sciences of the United States of America, 2001. **98**(9): p. 4899-903.
58. Smith, K.D., et al., *Structural basis of ligand binding by a c-di-GMP riboswitch.* Nature structural & molecular biology, 2009. **16**(12): p. 1218-23.
59. Ban, N., et al., *The complete atomic structure of the large ribosomal subunit at 2.4 A resolution.* Science, 2000. **289**(5481): p. 905-20.
60. Huang, L., J. Wang, and D.M. Lilley, *A critical base pair in k-turns determines the conformational class adopted, and correlates with biological function.* Nucleic acids research, 2016. **44**(11): p. 5390-8.
61. Huang, L. and D.M. Lilley, *The molecular recognition of kink-turn structure by the L7Ae class of proteins.* RNA, 2013. **19**(12): p. 1703-10.
62. Goody, T.A., et al., *The kink-turn motif in RNA is dimorphic, and metal ion-dependent.* RNA, 2004. **10**(2): p. 254-64.

63. Matsumura, S., Y. Ikawa, and T. Inoue, *Biochemical characterization of the kink-turn RNA motif.* Nucleic acids research, 2003. **31**(19): p. 5544-51.

64. Forster, T., **Zwischenmolekulare Energiewanderung Und Fluoreszenz.* Annalen Der Physik, 1948. **2**(1-2): p. 55-75.

65. Norman, D.G., et al., *Location of cyanine-3 on double-stranded DNA: Importance for fluorescence resonance energy transfer studies.* Biochemistry, 2000. **39**(21): p. 6317-6324.

66. McPhee, S.A., L. Huang, and D.M. Lilley, *A critical base pair in k-turns that confers folding characteristics and correlates with biological function.* Nature communications, 2014. **5**: p. 5127.

67. Lilley, D.M., *Analysis of branched nucleic acid structure using comparative gel electrophoresis.* Quarterly reviews of biophysics, 2008. **41**(1): p. 1-39.

68. Schroeder, K.T., P. Daldrop, and D.M. Lilley, *RNA tertiary interactions in a riboswitch stabilize the structure of a kink turn.* Structure, 2011. **19**(9): p. 1233-40.

69. Heppell, B. and D.A. Lafontaine, *Folding of the SAM aptamer is determined by the formation of a K-turn-dependent pseudoknot.* Biochemistry, 2008. **47**(6): p. 1490-9.

70. Turner, B., et al., *Induced fit of RNA on binding the L7Ae protein to the kink-turn motif.* RNA, 2005. **11**(8): p. 1192-200.

71. Koonin, E.V., P. Bork, and C. Sander, *A novel RNA-binding motif in omnipotent suppressors of translation termination, ribosomal proteins and a ribosome modification enzyme?* Nucleic acids research, 1994. **22**(11): p. 2166-7.

72. Nottrott, S., et al., *Functional interaction of a novel 15.5kD [U4/U6.U5] tri-snRNP protein with the 5' stem-loop of U4 snRNA.* The EMBO journal, 1999. **18**(21): p. 6119-33.

73. Marmier-Gourrier, N., et al., *A structural, phylogenetic, and functional study of 15.5-kD/Snu13 protein binding on U3 small nucleolar RNA.* RNA, 2003. **9**(7): p. 821-38.

74. Baird, N.J., et al., *YbxF and YlxQ are bacterial homologs of L7Ae and bind K-turns but not K-loops.* RNA, 2012. **18**(4): p. 759-70.

75. Cho, I.M., et al., *Ribosomal protein L7Ae is a subunit of archaeal RNase P.* Proceedings of the National Academy of Sciences of the United States of America, 2010. **107**(33): p. 14573-8.
76. Tycowski, K.T., et al., *A small nucleolar RNA requirement for site-specific ribose methylation of rRNA in Xenopus.* Proceedings of the National Academy of Sciences of the United States of America, 1996. **93**(25): p. 14480-5.
77. Tran, E.J., X. Zhang, and E.S. Maxwell, *Efficient RNA 2'-O-methylation requires juxtaposed and symmetrically assembled archaeal box C/D and C'/D' RNPs.* The EMBO journal, 2003. **22**(15): p. 3930-40.
78. Bleichert, F., et al., *A dimeric structure for archaeal box C/D small ribonucleoproteins.* Science, 2009. **325**(5946): p. 1384-7.
79. Ye, K., et al., *Structural organization of box C/D RNA-guided RNA methyltransferase.* Proceedings of the National Academy of Sciences of the United States of America, 2009. **106**(33): p. 13808-13.
80. Xue, S., et al., *Structural basis for substrate placement by an archaeal box C/D ribonucleoprotein particle.* Molecular cell, 2010. **39**(6): p. 939-49.
81. Lin, J., et al., *Structural basis for site-specific ribose methylation by box C/D RNA protein complexes.* Nature, 2011. **469**(7331): p. 559-63.
82. Omer, A.D., et al., *In vitro reconstitution and activity of a C/D box methylation guide ribonucleoprotein complex.* Proceedings of the National Academy of Sciences of the United States of America, 2002. **99**(8): p. 5289-94.
83. Watkins, N.J., A. Dickmanns, and R. Luhrmann, *Conserved stem II of the box C/D motif is essential for nucleolar localization and is required, along with the 15.5K protein, for the hierarchical assembly of the box C/D snoRNP.* Molecular and cellular biology, 2002. **22**(23): p. 8342-52.
84. Schultz, A., et al., *Protein-protein and protein-RNA contacts both contribute to the 15.5K-mediated assembly of the U4/U6 snRNP and the box C/D snoRNPs.* Molecular and cellular biology, 2006. **26**(13): p. 5146-54.

85. McKeegan, K.S., et al., *A dynamic scaffold of pre-snoRNP factors facilitates human box C/D snoRNP assembly.* Molecular and cellular biology, 2007. **27**(19): p. 6782-93.
86. Cantara, W.A., et al., *The RNA Modification Database, RNAMDB: 2011 update.* Nucleic acids research, 2011. **39**(Database issue): p. D195-201.
87. Schibler, U., D.E. Kelley, and R.P. Perry, *Comparison of methylated sequences in messenger RNA and heterogeneous nuclear RNA from mouse L cells.* Journal of molecular biology, 1977. **115**(4): p. 695-714.
88. Dominissini, D., et al., *Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq.* Nature, 2012. **485**(7397): p. 201-6.
89. Meyer, K.D., et al., *Comprehensive analysis of mRNA methylation reveals enrichment in 3' UTRs and near stop codons.* Cell, 2012. **149**(7): p. 1635-46.
90. Saneyoshi, M., F. Harada, and S. Nishimura, *Isolation and characterization of N6-methyladenosine from Escherichia coli valine transfer RNA.* Biochimica et biophysica acta, 1969. **190**(2): p. 264-73.
91. Iwanami, Y. and G.M. Brown, *Methylated bases of ribosomal ribonucleic acid from HeLa cells.* Archives of biochemistry and biophysics, 1968. **126**(1): p. 8-15.
92. Dimock, K. and C.M. Stoltzfus, *Sequence specificity of internal methylation in B77 avian sarcoma virus RNA subunits.* Biochemistry, 1977. **16**(3): p. 471-8.
93. Beemon, K. and J. Keith, *Localization of N6-methyladenosine in the Rous sarcoma virus genome.* Journal of molecular biology, 1977. **113**(1): p. 165-79.
94. Bringmann, P. and R. Luhrmann, *Antibodies specific for N6-methyladenosine react with intact snRNPs U2 and U4/U6.* FEBS letters, 1987. **213**(2): p. 309-15.
95. Bokar, J.A., et al., *Purification and cDNA cloning of the AdoMet-binding subunit of the human mRNA (N6-adenosine)-methyltransferase.* RNA, 1997. **3**(11): p. 1233-47.

96. Niu, Y., et al., *N6-methyl-adenosine (m6A) in RNA: an old modification with a novel epigenetic function.* Genomics, proteomics & bioinformatics, 2013. **11**(1): p. 8-17.
97. Wang, X., et al., *N6-methyladenosine-dependent regulation of messenger RNA stability.* Nature, 2014. **505**(7481): p. 117-20.
98. Wei, C.M. and B. Moss, *Nucleotide sequences at the N6-methyladenosine sites of HeLa cell messenger ribonucleic acid.* Biochemistry, 1977. **16**(8): p. 1672-6.
99. Jia, G., et al., *N6-methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO.* Nature chemical biology, 2011. **7**(12): p. 885-7.
100. Zheng, G., et al., *ALKBH5 is a mammalian RNA demethylase that impacts RNA metabolism and mouse fertility.* Molecular cell, 2013. **49**(1): p. 18-29.
101. Gerken, T., et al., *The obesity-associated FTO gene encodes a 2-oxoglutarate-dependent nucleic acid demethylase.* Science, 2007. **318**(5855): p. 1469-72.
102. Zhao, X., et al., *FTO-dependent demethylation of N6-methyladenosine regulates mRNA splicing and is required for adipogenesis.* Cell research, 2014. **24**(12): p. 1403-19.
103. Jia, G., Y. Fu, and C. He, *Reversible RNA adenosine methylation in biological regulation.* Trends in genetics : TIG, 2013. **29**(2): p. 108-15.
104. Meyer, K.D., et al., *5′ UTR m(6)A Promotes Cap-Independent Translation.* Cell, 2015. **163**(4): p. 999-1010.
105. Wang, X., et al., *N(6)-methyladenosine Modulates Messenger RNA Translation Efficiency.* Cell, 2015. **161**(6): p. 1388-99.
106. Beaucage, S.L. and M.H. Caruthers, *Deoxynucleoside Phosphoramidites - a New Class of Key Intermediates for Deoxypolynucleotide Synthesis.* Tetrahedron Letters, 1981. **22**(20): p. 1859-1862.
107. Hakimelahi, G.H., Z.A. Proba, and K.K. Ogilvie, *High-Yield Selective 3'-Silylation of Ribonucleosides.* Tetrahedron Letters, 1981. **22**(52): p. 5243-5246.

108. Bassi, G.S., et al., *Ion-induced folding of the hammerhead ribozyme: a fluorescence resonance energy transfer study.* The EMBO journal, 1997. **16**(24): p. 7481-9.
109. Clegg, R.M., *Fluorescence Resonance Energy-Transfer and Nucleic-Acids.* Methods in enzymology, 1992. **211**: p. 353-388.
110. Clegg, R.M., *Fluorescence resonance energy transfer and nucleic acids.* Methods in enzymology, 1992. **211**: p. 353-88.
111. Bhattacharyya, A., A.I. Murchie, and D.M. Lilley, *RNA bulges and the helical periodicity of double-stranded RNA.* Nature, 1990. **343**(6257): p. 484-7.
112. Ashraf, S., L. Huang, and D. Lilley, *Sequence determinants of the folding properties of box C/D kink-turns in RNA.* RNA, 2017.

8. List of Abbreviations

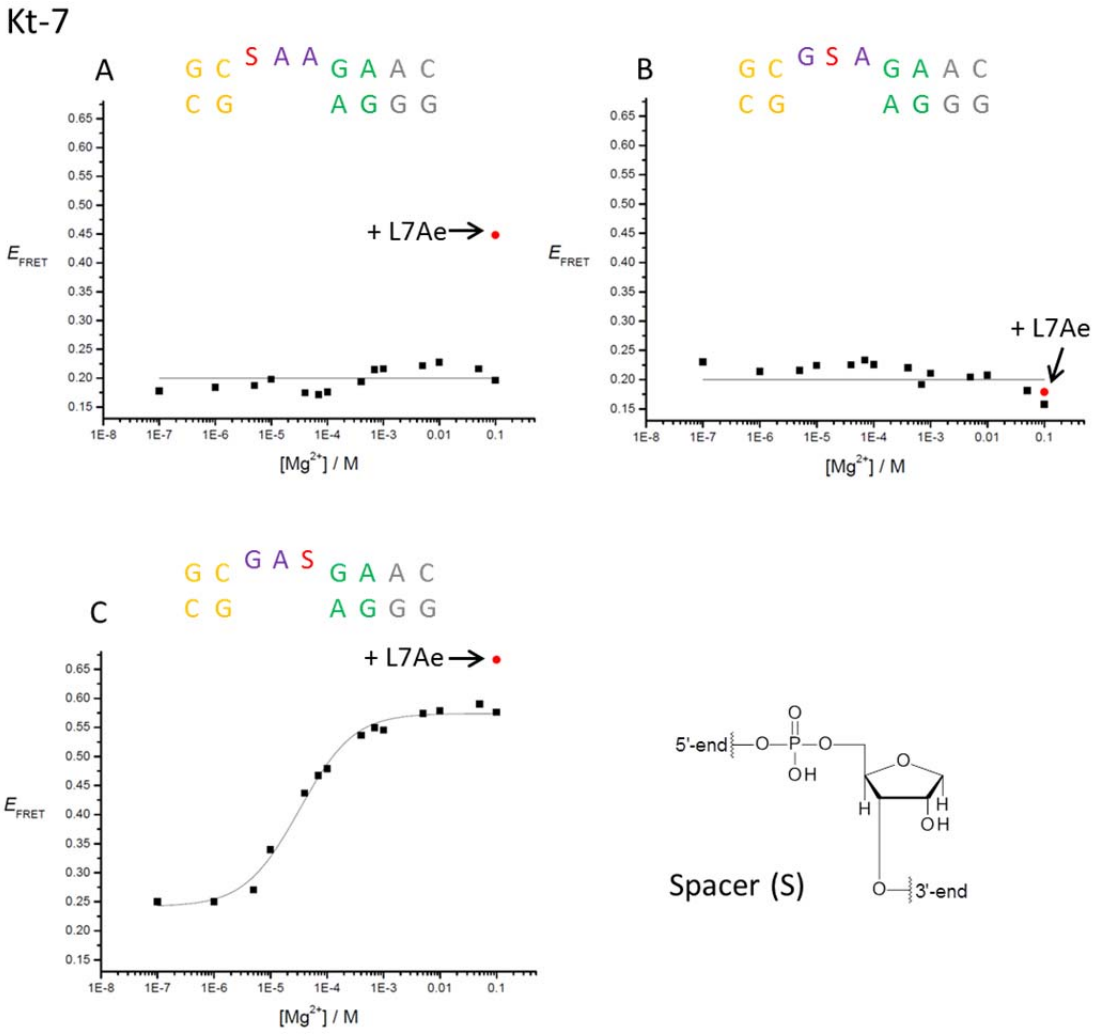
%	percent
˚C	degree Celsius
1X TB	90mM Tris.borate
1X TBE	90mM Tris.borate and 10mM EDTA buffer
A	adenine
A_{490}	acceptor absorbance at 490nm wavelength
A_{547}	acceptor absorbance at 547nm wavelength
Ac	acetyl
Af	*Archeoglobus fulgidus*
APS	ammonium persulfate
C	cytosine
C helix	canonical helix
CE	cyanoethyl
CGE	comparative gel electrophoresis
CPG	controlled pore glass
D_{490}	donor absorbance at 490nm wavelength
dmf	dimethylformamidino
DMSO	dimethyl sulfoxide

DNA	deoxyribonucleic acid
EDTA	ethylenediaminetetraacetic acid
E_{FRET}	efficiency of fluorescence resonance energy transfer
ETT	5-Ethylthio-1H-tetrazole
FRET	fluorescence resonance energy transfer
G	guanine
GMP	guanosine monophosphate
h	hour
h^{-1}	per hour
HCl	hydrochloric acid
Hm	*Haloarcula marismortui*
Hs	*Homo sapiens*
HPLC	high pressure liquid chromatography
I	inosine
iPr	isopropyl
ITC	isothermal titration calorimetry
lncRNA	long non-coding RNA
M	molar
m6A	N6-methyl adenine
Mg^{2+}	magnesium ion

$MgCl_2$	magnesium chloride
mins	minutes
miRNA	micro RNA
mL	millilitre
mM	millimolar
mm	millimetre
mRNA	messenger RNA
NaCl	sodium chloride
NC helix	non-canonical helix
nm	nanometre
nM	nanomolar
Pac	phenoxyacetyl
pmol	picomoles
RNA	ribonucleic acid
RNP	ribonucleoprotein
RT	room temperature
SAM	*S*-adenosyl methionine
siRNA/RNP	small interfering RNA/RNP
snoRNA/RNP	small nucleolar RNA/RNP
snRNA/RNP	small nuclear RNA/RNP

sRNA	small RNA
T. thermophilus	*Thermus thermphilus*
Tac	tert-butylphenoxyacetyl
t-BDMS	*tert*-butyldimethylsilyl
TCA	trichloroacetic acid
TEA	triethylamine
TEA, 3HF	triethylamine trihydrofluoride
TEMED	tetramethylethylenediamine
tRNA	transfer RNA
U	uracil
UTRs	untranslated regions
V	volts
W	Watts
µg	microgram
µL	microlitre
µM	micromolar
µmol	micromole

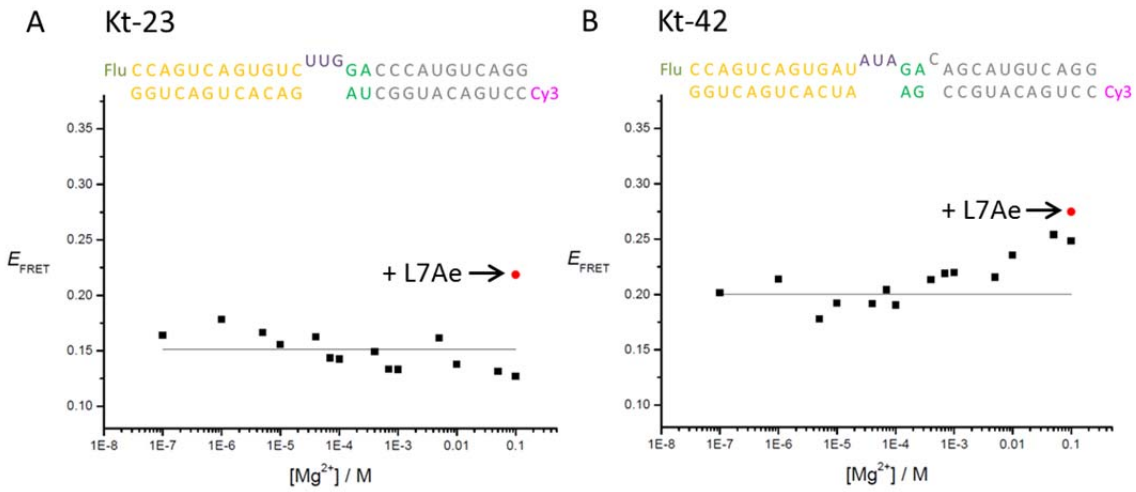
9. Appendix



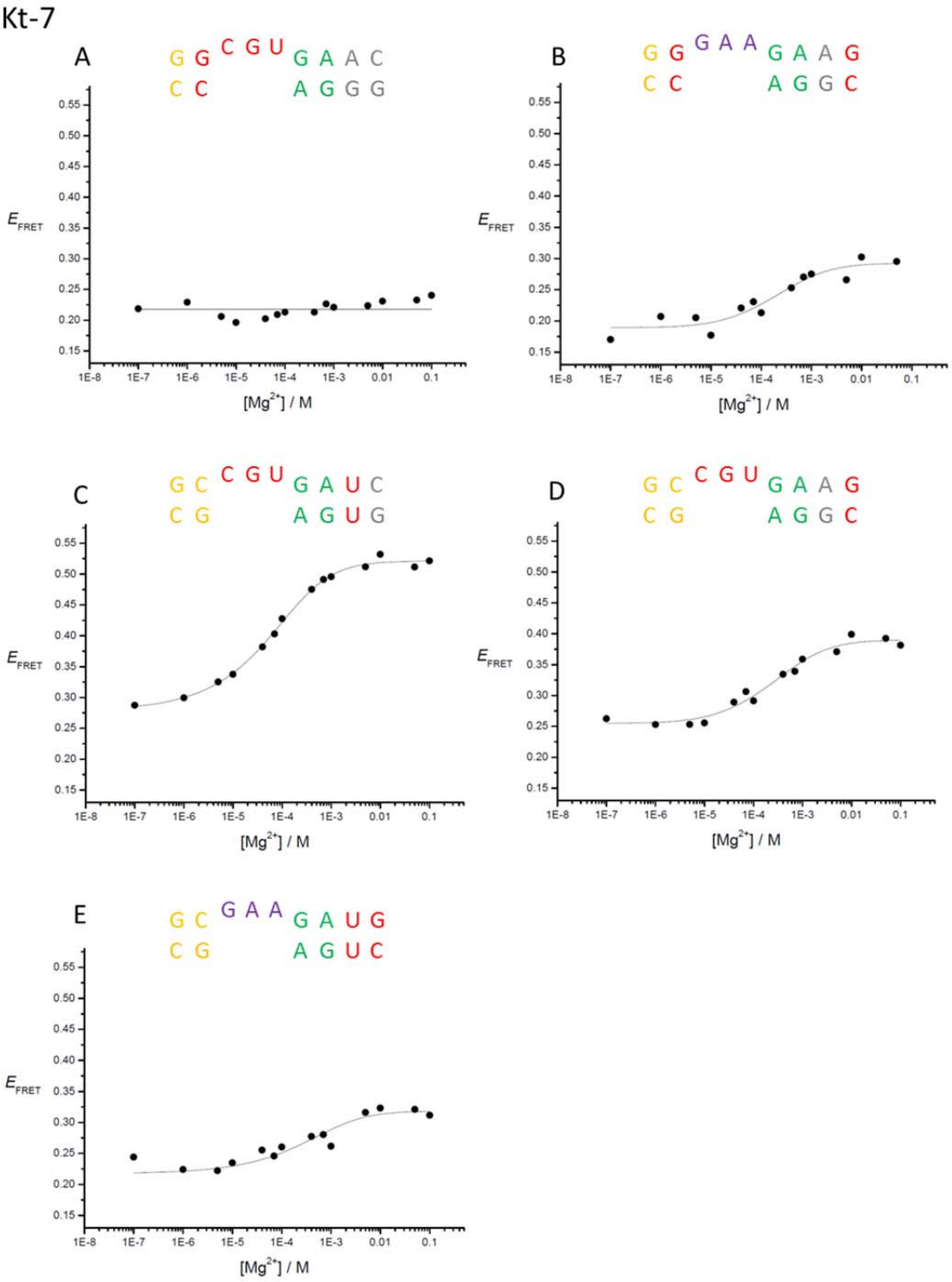
Appendix Figure 1: Mg^{2+} ion-dependent folding analysis by FRET of *Hm*Kt-7 variants modified by the individual removal of nucleobases from the loop positions. The sequence of the core k-turn region is displayed at the top of each diagram with the modification introduced highlighted in red. S = Spacer, a nucleotide without any nucleobase, the structure of which is displayed on the bottom right. E_{FRET} (y-axis) has been plotted as a function of Mg^{2+} ion concentration (x-axis) and the data fitted to the two-state model for A. L1:S, B. L2:S and C. L3:S modified Kt-7 variants. After the magnesium ion titrations, 1µM L7Ae was added and E_{FRET} calculated which is shown as the red circle.

3b \ 3n	A	C	G	U
A	0.24	0.35	0.57	0.29
C	0.54	0.57	0.22	0.52
G	0.18	0.18	0.51	0.22
U	0.30	0.16	0.53	0.42

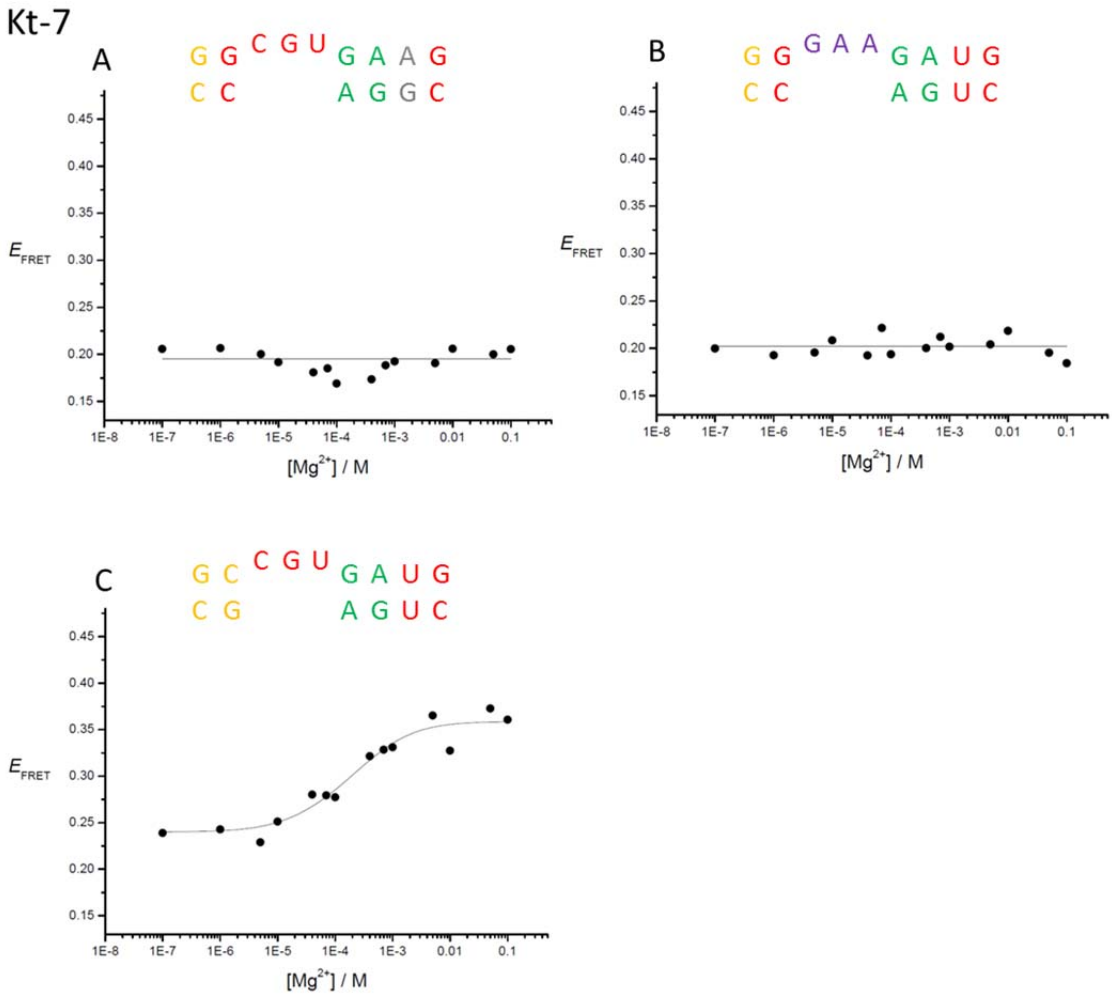
Appendix Figure 2: The 3b:3n grid depicting the final E_{FRET} values for all the 3b:3n modified *Hm*Kt-7 variants analysed by FRET for folding properties in Mg^{2+} ions. The 3b:3n sequences that fold well in Mg^{2+} ions ($E_{FRET} \geq 0.5$) are shaded red, while those with intermediate folding ability are displayed as white, and ones which are unable to fold ($E_{FRET} \leq 0.3$) are shaded blue.



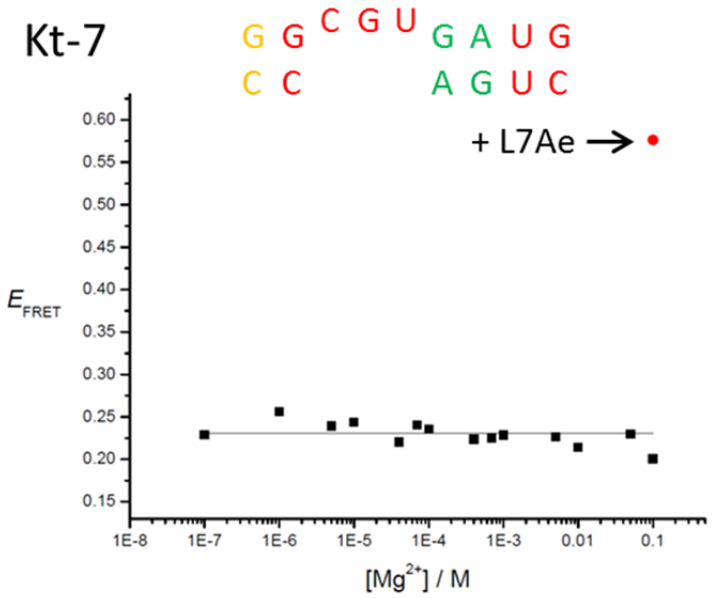
Appendix Figure 3: Mg^{2+} ion-dependent folding analysis of wild-type *Hs*Kt-23 and *Hs*Kt-42 by FRET. The sequences of the wild-type Kt-23 (A) and Kt-42 (B) constructs used in the FRET experiments are displayed at the top. E_{FRET} (y-axis) has been plotted as a function of Mg^{2+} ion concentration (x-axis) and the data fitted to the two-state model for A. wild-type Kt-23 and B. wild-type Kt-42. After the magnesium ion titrations, 1μM L7Ae was added and E_{FRET} calculated which is shown as the red circle.



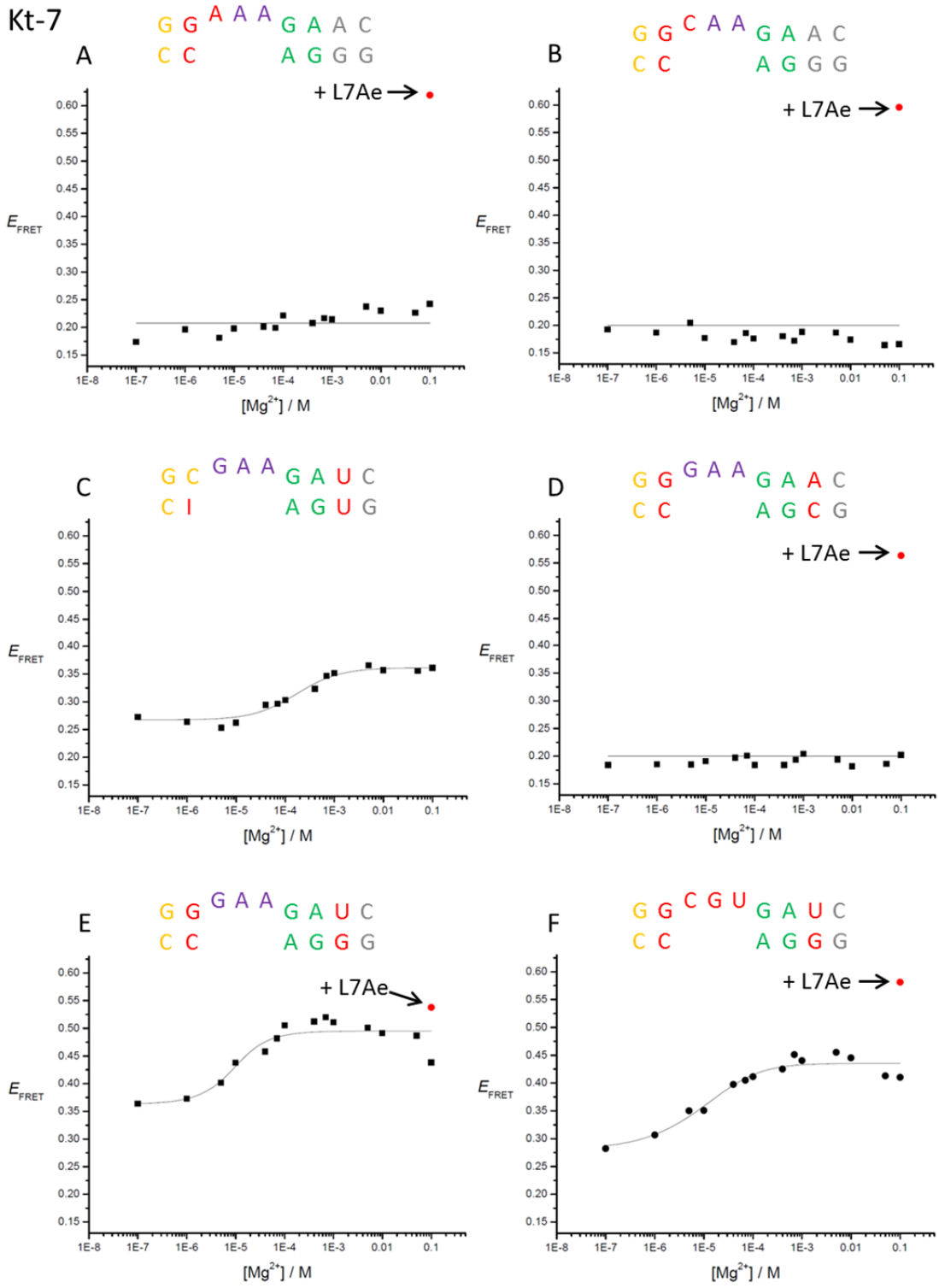
Appendix Figure 4: Mg^{2+} ion-dependent folding analysis by FRET of *Hm*Kt-7 variants doubly modified with *Af*boxC/D sequence elements. The sequence of the core k-turn region is displayed at the top of each diagram with the modifications introduced highlighted in red. E_{FRET} (y-axis) has been plotted as a function of Mg^{2+} ion concentration (x-axis) and the data fitted to the two-state model for A. -1GC + LCGU, B. -1GC + 4GC, C. LCGU +3UU, D. LCGU + 4GC, and E. 3UU + 4GC modified Kt-7 variants.



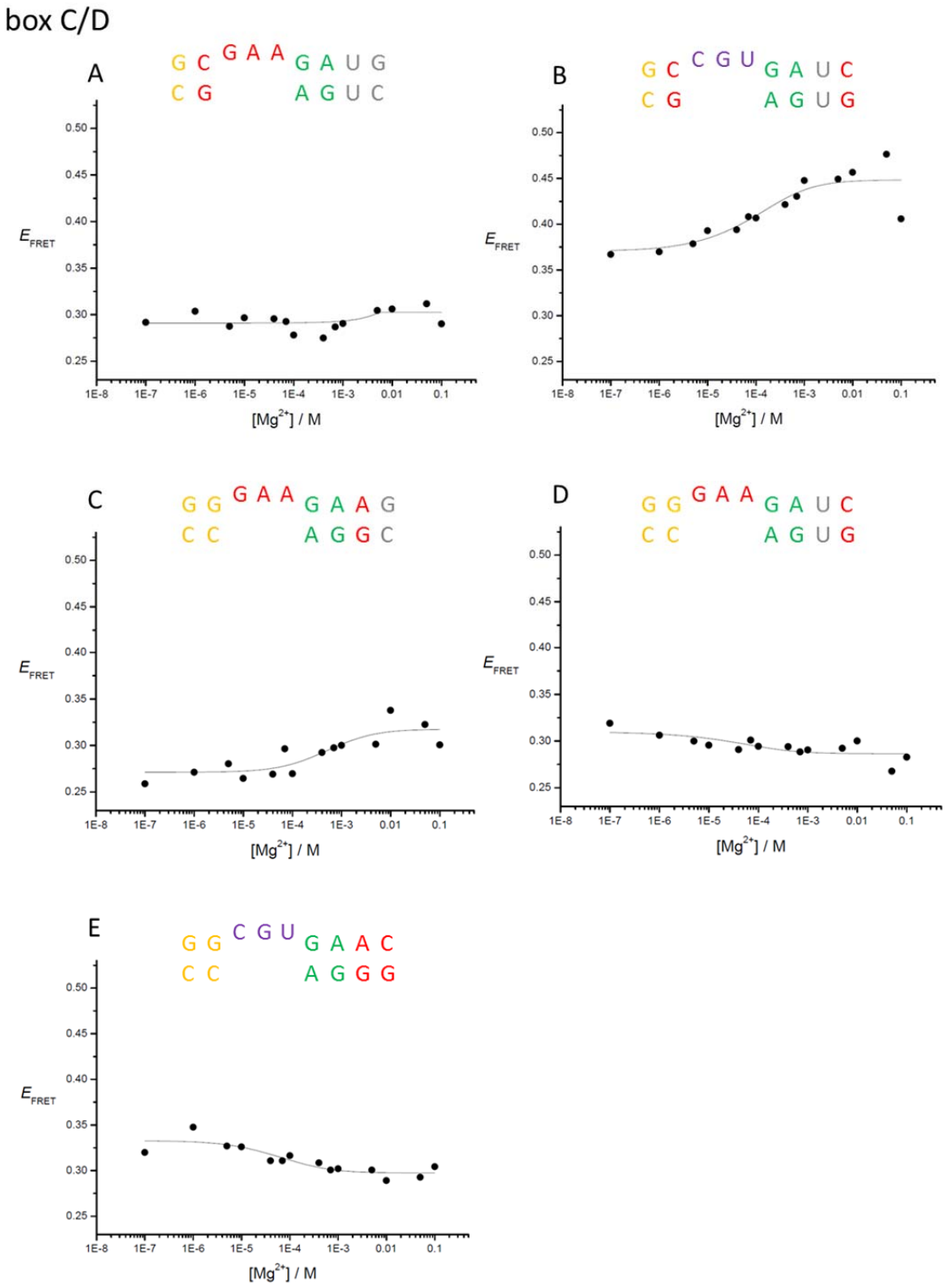
Appendix Figure 5: Mg^{2+} ion-dependent folding analysis by FRET of *Hm*Kt-7 variants triply modified with *Af*boxC/D sequence elements. The sequence of the core k-turn region is displayed at the top of each diagram with the modifications introduced highlighted in red. E_{FRET} (y-axis) has been plotted as a function of Mg^{2+} ion concentration (x-axis) and the data fitted to the two-state model for A. -1GC + LCGU + 4GC, B. -1GC + 3UU + 4GC and C. LCGU +3UU + 4GC modified Kt-7 variants.



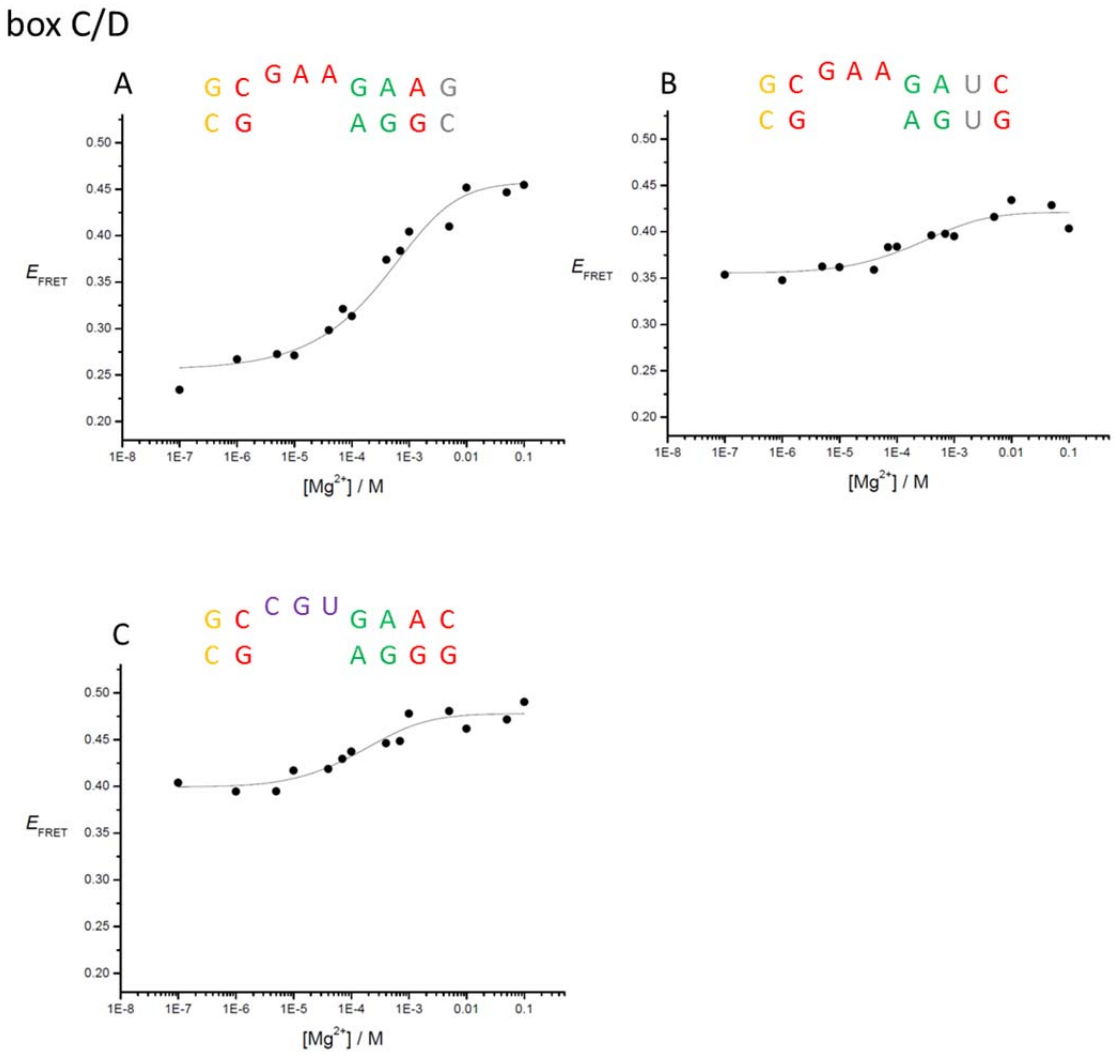
Appendix Figure 6: Mg^{2+} ion-dependent folding analysis by FRET of *Hm*Kt-7 modified with all four *Af*boxC/D sequence elements. The sequence of the core k-turn region is displayed at the top with the modifications introduced highlighted in red. E_{FRET} (y-axis) has been plotted as a function of Mg^{2+} ion concentration (x-axis) and the data fitted to the two-state model for the -1GC + LCGU + 3UU + 4GC modified Kt-7 variant. After the magnesium ion titrations, 1μM L7Ae was added and E_{FRET} calculated which is shown as the red circle.



Appendix Figure 7: Mg^{2+} ion-dependent folding analysis of modified *Hm*Kt-7 variants by FRET. The sequence of the core k-turn region is displayed at the top of each diagram with the modifications introduced highlighted in red. E_{FRET} (y-axis) has been plotted as a function of Mg^{2+} ion concentration (x-axis) and the data fitted to the two-state model for A. -1GC + L1:A, B. -1GC + L1:C, C. -1n:I +3UU, D. -1GC + 3AC, E. -1GC + 3UG and F. -1GC + LCGU + 3UG modified Kt-7 variants. In some cases, after the magnesium ion titrations, 1μM L7Ae was added and E_{FRET} calculated which is shown as the red circle.



Appendix Figure 8: Mg^{2+} ion-dependent folding analysis by FRET of *Af*boxC/D variants doubly modified with *Hm*Kt-7 sequence elements. The sequence of the core k-turn region is displayed at the top of each diagram with the modifications introduced highlighted in red. E_{FRET} (y-axis) has been plotted as a function of Mg^{2+} ion concentration (x-axis) and the data fitted to the two-state model for A. -1CG + LGAA, B. -1CG + 4CG, C. LGAA +3AG, D. LGAA + 4CG, and E. 3AG + 4CG modified boxC/D variants.



Appendix Figure 9: Mg^{2+} ion-dependent folding analysis by FRET of *Af*boxC/D variants triply modified with *Hm*Kt-7 sequence elements. The sequence of the core k-turn region is displayed at the top of each diagram with the modifications introduced highlighted in red. E_{FRET} (y-axis) has been plotted as a function of Mg^{2+} ion concentration (x-axis) and the data fitted to the two-state model for A. -1CG + LGAA + 3AG, B. -1CG + LGAA + 4CG and C. -1CG + 3AG + 4CG modified boxC/D variants.